Université Libre De Bruxelles

Faculté des Sciences Appliquées

Théorie de l'Information et des Communications

Communication avec des variables quantiques continues: clonage et cryptographie

Thèse présentée en vue de l'obtention
du grade de Docteur en Sciences Appliquées

Sofyan Iblisdir

Promoteur: Nicolas Cerf

ANNÉE ACADEMIQUE : 2002–2003

Contents

Remerciements

Je tiens avant tout à exprimer ma gratitude au Professeur Nicolas Cerf. Je le remercie de m'avoir fait découvrir la théorie quantique de l'information, et de m'avoir communiqué son enthousiasme pour cette discipline en me proposant cette thèse. Je le remercie aussi pour tout ce qu'il ma appris. Sa guidance, ses encouragements, et l'environnement propice à la recherche qu'il a su créer ont été mon premier soutien. Je remercie aussi le Professeur Serge Massar auprès de qui j'ai aussi beaucoup appris durant ces quatre années. Je remercie les autres personnes avec lesquelles j'ai eu la chance de collaborer dans le cadre de cette thèse: Peter van Loock, Gilles Van Assche et Jaromir Fiurasek. Je remercie aussi Claude Archer, Anne-Cécile Muffat, Patrick Navez, Frédéric Grosshans, Stefano Pironio, Jonathan Barrett et Louis Lamoureux pour leur soutien, en particulier durant la phase finale cette thèse. Enfin, ma gratitude va à mes amis, à Irina, à ma famille, et en particulier à ma mère.

Résumé

Un problème fondamental posé par la théorie quantique de l'information est de déterminer comment l'information contenue dans l'état d'un système quantique peut être répartie dans plusieurs autres systèmes. Deux sujets où ce problème joue une rôle central sont ici étudiés. Le clonage et la cryptographie quantique. Ces deux sujets sont par ailleurs reliés. Nous considérons principalement le cas de variables quantiques continues. Nous justifions cet intérêt par l'importance croissante prise par ce type de système dans les communications quantiques.

Dans la première partie de cette thèse, nous nous intéressons au clonage de systèmes discrets et surtout continus. Nous établissons des bornes sur le clonage optimal, puis proposons des machines atteignant ces bornes. Cette étude permet de comparer diverses manières d'encoder de l'information dans un système quantique.

Ensuite, nous nous intéressons à la cryptographie quantique, et étudions la sécurité de protocoles continus de distribution quantique de clef. En particulier, nous analysons la sécurité de protocoles prometteurs d'un point de vue expérimental: les protocoles à états cohérents, et montrons comment ces protocoles peuvent être rendus sûrs et efficaces.

Chapter 1

Introduction

Information and quantum mechanics

Since the seminal work of Shannon [1], we have an operational understanding of information well adapted to study problems of communication. Information is a quantity measuring our ignorance of the outcome of a statistical experiment[1]. Interestingly, there is a sense in which this notion of information matches our familiar conceptions. For example, what we find to be (relevant) information when listening to the radio news is precisely the part of what the speaker says that we couldn't foresee.

The applications of the mathematical theory of information are countless. They can be found in (tele-)communications, of course, or Physics, but also in Biology or Linguistics for example [2]. The main feature of information theory making it powerful is that it deals only with abstract objects (probability distribution of random variables), and give us so-called *coding theorems*, ruling the way information can be efficiently and reliably transmitted. But for this very reason, the theory is also limited because it leaves aside a crucial feature of information: it is always represented by the state of a physical system. That is, the statistical experiments with which information is concerned are physical experiments. The major achievement of quantum information theory is to have shown that Shannon's theory is in fact limited to those experiments involving classical systems. When information is represented by quantum systems, one can go beyond this theory and define a new kind of information: quantum information [3].

> "Quantum information is that kind of information which is carried by quantum systems from a preparation device to a measuring apparatus in a quantum mechanical experiment."

This kind of information is distinct from classical information. Indeed, classical information, the kind of information described by Shannon's theory, is fungible. It can be, at least in principle, transparently converted from one representation to another. We can, for example, read a text on a sheet of paper, and orally transmit it to someone who will store it in the form of bits on the hard disk of a computer without the information undergoing any damage during these operations. Quantum information enjoys similar properties. For example, we can, in principle, transfer the quantum information encoded in the polarisation state of a photon into the spin state of an electron, etc. But quantum information cannot be converted into any equivalent (representation of) classical information. That is, it cannot be transformed into classical

[1] We are not concerned here in the concept of information as opposed to "misinformation". The context of this thesis is purely technical.

information in such a way that it can be transformed back into quantum information transparently: the device depicted in Fig.1.1, which we call a classical teleporter is an impossible machine. Indeed, since the classical information in between is accessible, a classical teleporter would allow us to acquire knowledge about a quantum state without disturbing it.

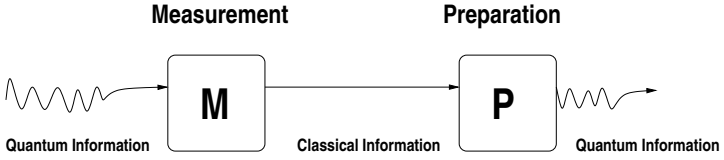

Figure 1.1: Division of a torus or a sphere into four regions.

We could present quantum information theory as founded on a fundamental questioning of classical information theory. But this wouldn't reflect how the field developed. Rather, quantum information emerged from the discovery and the study of quite a small number of brilliant applications where the quantum character of the systems used to perform these tasks plays a crucial role. We will here briefly survey these applications.

At first sight, we might wonder what can we gain from using quantum instead of classical carriers of information, and whether there is any advantage in considering quantum information. Indeed, when we think of the principles of quantum mechanics, we usually have in mind statements such as:

- It is in general impossible to assign simultaneously definite values to non-commuting observable.
- One cannot (in general) deterministically predict the result of a quantum measurement.
- One cannot (in general) measure a quantum system, that is, get information about its state, without disturbing it.

The laws of quantum mechanics are often perceived negatively, and indeed they certainly limit our ability to manipulate quantum systems. But, as will be shown now, quantum behavior can also turn out to be a resource.

Quantum computation

Quantum computers. Just as the fact of using quantum carriers has led to extend our conception of information, the possibility to perform a different kind of computation, quantum computation, has been considered. A quantum computer is a device that would operate on arrays of "quantum bits" (or qubits), that is two-level quantum systems, in much the same way as a classical computer operates on arrays of classical bits, that is two-level classical systems. The interest of quantum computers lies in

the possibility of preparing *superpositions* of states of qubits. These superpositions would in some cases provide a natural means of performing parallel computation, and exponentially outperform classical computers. Let us briefly survey two possible uses of a quantum computer.

Efficient factoring. The best known quantum algorithm is probably Shor's algorithm for factorisation [4]. Factoring is an example of a problem where solutions are easy to verify once found, but hard to find. Let n denote an integer. Up to now, the best algorithms we have for extracting its prime factors from n using classical resources are superpolynomial in $\log n$. Specifically, the best known factoring algorithm (the "number field sieve") requires about

$$\exp(c(\ln n)^{1/3}(\ln \ln n)^{2/3}) \tag{1.1}$$

steps, where $c \approx 1.9$. Shor proved, by explicit construction, that with quantum carriers and using the superposition principle, one could factorise n in about

$$O((\ln n)^3) \tag{1.2}$$

steps. This is an *exponential* speedup. Apart from the implications of Shor's algorithm for computational complexity theory, it also has a practical impact, because the presumed difficulty of factorisation is the basis of many widely used public key cryptographic schemes, such as RSA [5].

Efficient simulation of quantum systems. Another possible use of a quantum computer is the simulation of the behaviour of quantum mechanical systems. Let us consider a simple example, and suppose that we want to simulate a system of N interacting two-level quantum systems. Such a system is described by $2^{N+1} - 2$ real parameters (if its state is pure), so that the amount of resources required to simulate it grows exponentially fast with its size. It is thus practically impossible to simulate a quantum mechanical system on a classical computer. On another hand, the elementary bit of memory of a quantum computer, the "qubit", would itself be a two-level quantum system, so that the amount of resources now only grows linearly with the size of the system to simulate. This fact alone is not sufficient to assert that quantum systems can always be efficiently simulated on a quantum computer, but at least, it leaves open such a possibility, at least in some cases. For example, promising research has shown how some quantum systems, such as a lattice of fermions, could be efficiently simulated on a quantum computer [6].

Quantum error correction

Quantum computers appear to be beautiful devices, but perhaps too beautiful, and one can wonder whether there isn't a price to pay for the exponential advantage they offer over classical computers. From an experimental point of view, it is a very challenging task to prepare quantum states on demand and to manipulate them at will. Still, preparing and manipulating quantum information is not the biggest problem, and actually very encouraging progress has been achieved during the last five to ten years (see [7] and references therein). The real issue is noise. Information encoded in quantum systems is very delicate, and one can wonder whether it is possible to protect it against the unavoidable noise coming from the system-environment interaction, the decoherence, and from the non-perfection of quantum logical gates. If yes, then would the amount of necessary resources grow dramatically with the size of the state to protect, thus annihilating the benefit gained from quantum computation? These questions have been answered by the discovery of quantum error correcting codes (and fault-tolerant quantum computation).

As their classical counterpart, quantum codes allow the protection of quantum information so as to make it resilient to noise. The most important result of the theory of quantum error correcting codes for quantum computation are the threshold theorems [7]. Loosely speaking, these theorems tell us that if quantum operations can be performed sufficiently accurately, that is with a (properly quantified) accuracy above a certain threshold, then quantum error correcting codes can be efficiently used to achieve quantum computation that is arbitrarily close to perfect.

Historically speaking, the discovery of quantum error correcting codes marks the second boost of interest for quantum information theory and quantum computation [2]. And indeed, this result is crucial because it shows that the difficulties encountered towards achieving quantum computation are essentially of technological order. They are not rooted in any fundamental physical principle.

Quantum Communication

Quantum teleportation. Along with quantum computation, it is certainly worth saying a few words about quantum teleportation, even if we will not talk about it again in the remainder of this thesis. Quantum teleportation perfectly illustrates the distinction between classical and quantum information, and beautifully demonstrates that the features of quantum mechanics, in particular non-local correlations, or quantum entanglement, are not only weird properties of nature at the quantum level, but also valuable resources. In this sense, quantum teleportation, more appropriately called entanglement-assisted teleportation, is the paradigm of quantum information theory.

We have seen that it is impossible to teleport classically a quantum state. Quite surprisingly, when supplemented with non-local resources, teleportation becomes a feasible task . We will not need the precise description of entanglement-assisted teleportation here [8]. But Fig.1 suffices to state our point. Thanks to entanglement, the quantum information is *non-locally* transfered from A to B. In contrast to classical teleportation, the in-between classical information now never contains any information about the teleported quantum information, it only conveys information on how to recover the quantum information.

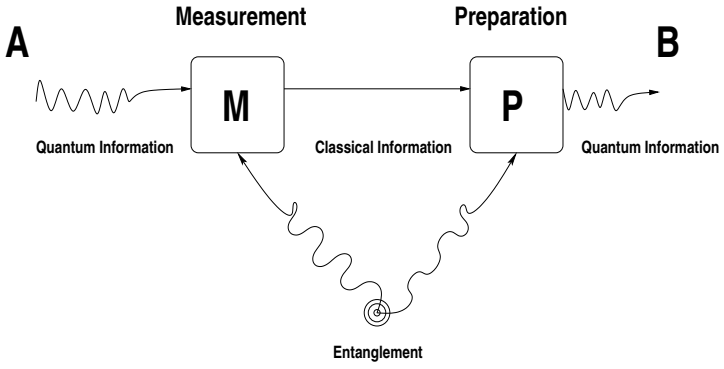

Figure 1.2: Entanglement-assisted teleportation

Quantum cryptography. Quantum cryptography, more properly called quantum key distribution, aims at providing two remote parties with a secure means to get

[2]The first boost came up with the publication of Shor's algorithm for factorisation.

a common secret bit string, the key. This key can later be used for confidential communication. The fundamental idea in quantum cryptography is to use non-orthogonal quantum states to transmit the key, in such a way that any intervention of a potential eavesdropper will cause a detectable disturbance.

Quantum cryptography plays a central role in quantum information. From an experimental point of view, it has definitely proven its feasibility [9, 10]. From a theoretical point of view, the study of quantum cryptography is very much related to other branches of quantum information theory, such as the theory of quantum error correcting codes. This will be explained in details in chapters 5 and 6, which are entirely devoted to quantum cryptography.

Experimental issues

To determine whether we will someday be able to build a quantum computer is a very controversial matter. Many eminent scientists are very doubtful about that (see [11] for example), and they might be right. However, one can wonder whether the intermediate research towards building a quantum computer is worth undertaking even if we don't eventually get a quantum computer? Looking at the past 10 years, we see that the quest for a reliable implementation of a quantum computer has fostered an unprecedented interest in ion traps, NMR engineering or Josephson junctions [7]. This line of research directly enhance our ability to *control* quantum systems. My opinion is that this ability will (very) soon play a crucial role in both fundamental and applied Physics, even if not to build a quantum computer.

Unlike quantum computation, quantum communication *does* lie in the scope of current technology, because the typical size of the quantum systems to manipulate here is much smaller than for quantum computation, and because much simpler manipulations are made on quantum information carriers. In most situations, we only require the ability to prepare and measure a small number of quantum systems, but we don't need to make them interact. Therefore, the use of light modes has proven a reliable way to perform quantum communication tasks such as quantum cryptography or quantum teleportation. Truly convincing experiments have been performed. Just to give an example, it has been possible to perform quantum teleportation with photons over distances of about 2 km [12]. In turn, the relative ease in performing quantum information processing with light has driven research to focus mostly on that kind of implementation. This is why, in this thesis, we will implicitly (or explictly) restrict to the case where quantum information is carried by light modes.

Content of this thesis and main results

This thesis is devoted to two subjects: quantum cloning and quantum key distribution. The first way in which these two subjects are related is that both are concerned with the same question, fundamental in quantum information: how does the information contained in a quantum system distribute amongst several quantum systems? The second way in which these two subjects are related is that the formalism of quantum cloning provides a means to study the security of quantum key distribution protocols.

We will mostly consider quantum continuous variables. These quantum information carriers have emerged during the last five years as a promising alternative to discrete quantum variables. Continuous variable quantum cryptographic schemes, for instance, seem to allow for facilitated implementations and higher secret key generation rates than their discrete counterparts which often require difficult preparation and measurement of single-photon states representing two-level quantum systems.

This thesis is organised as follows. In chapter 2, we present the fundamentals of quantum information theory. This chapter introduces the concepts and results that will be explicitly and implicitly used in the subsequent chapters. In chapter 3, the issue of cloning to which this thesis is largely devoted is introduced and discussed. A special class of quantum cloning machines is then presented, where new evidence are brought that the informational content of a pair of orthogonal quantum states may be higher than that of a pair of identical states. Chapter 4 is a detailed analysis of continuous variable quantum cloning. The issues of cloning and phase-conjugation are considered, as well phase-conjugate input quantum cloning machines. The last two chapters are concerned with quantum key distribution. The principles of quantum key distribution (QKD) are reviewed in chapter 5 as well as continuous variable QKD schemes. Chapter 6 aims at analysing the security of a particular class of continuous variable protocols: the coherent-state protocols. We finally conclude in chapter 7 by discussing open questions and future lines of research.

The main results of this thesis as well as their publication status are presented in the following list.

- Upper bounds are given for optimal Gaussian cloning transformations turning N identical replicas onto M clones ($M \geq N$). Publication in Physical Review A (Rapid Communications) [13].
- Implementations achieving optimal cloning are proposed. Publication in Physical Review Letters [14].
- Phase conjugation of quantum continuous variables is studied. Publication in Physical Review A [15].
- Quantum cloning machines with phase-conjugate input modes have been studied. Publication in Physical Review Letters [16].
- Quantum cloning machines for orthogonal qubits have been proposed. Publication in Physical Review A [17].
- The security of coherent-state protocols under general conditions has been studied. Article in preparation (Joint work with Gilles Van Assche).

Chapter 2

Fundamentals of quantum information

We introduce the concepts, vocabulary and results that will be used in the next chapters. The following terms can be considered as "new" with respect to a standard book of quantum mechanics: qubits, positive operator valued measure, cp-map, entanglement, quantum coding.

2.1 Quantum Information Carriers

Let us start by recalling our definition of quantum information. Quantum Information is carried by quantum systems from a preparation device to a measuring apparatus in a quantum mechanical experiment [3]. Now a quantum system may have many degrees of freedom. Let us consider an example where our quantum information carrier is an isolated electron. The Hilbert space describing this system is the tensor product $\mathcal{H}_q \otimes \mathcal{H}_s$, where $\mathcal{H}_q = L^2(\mathbf{R}^3)$ is the Hilbert space associated with the position of this electron and $\mathcal{H}_s = \mathbf{C}^2$ is the Hilbert space associated with its spin degree of freedom. The state of this electron generally reads:

$$|\Phi\rangle = \sum_{s=\downarrow,\uparrow} \int d\mathbf{q}\ \phi(\mathbf{q}, s)\ |\mathbf{q}\rangle|s\rangle.$$

Now, if the preparation device encodes quantum information only in the spin degree of freedom of the electron, and that only the spin is measured by our measuring device, then we are only interested in the "substate"

$$\mathrm{Tr}_{\mathcal{H}_q}|\Phi\rangle\langle\Phi| = \sum_{s,s'} c_{s,s'}|s\rangle\langle s'|,$$

where $c_{s,s'} = \int\ d\mathbf{q}\psi(\mathbf{q}, s)\bar{\psi}(\mathbf{q}, s')$. This is a great simplification that is generally encountered in quantum information. Although the complete description of any physical system ultimately requires an infinite-dimensional Hilbert space, we can restrict to a smaller Hilbert space, which is often finite-dimensional.

The quantum information of our example is a quantum bit or qubit. The qubit can be thought of as the quantum analogue of a classical bit. It is the simplest non-trivial piece of quantum information: it is just the information carried by a two-level quantum system. Similarly, the information carried by a d-level quantum system is called a qudit, etc.

2.2 States and measurements

States

Let $\mathcal{H}$ and $\mathcal{B}(\mathcal{H})$ denote respectively the Hilbert space associated with a quantum system and the algebra of bounded operators on $\mathcal{H}$. It is well-known that the set of possible states (density operators) for this quantum system is

$$\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H}) |\ \rho \geq 0,\ \mathrm{Tr}\rho = 1\}$$

This set is closed under convex combinations, i.e.

$$\forall \rho_1, \rho_2 \in \mathcal{S}(\mathcal{H}), 0 \leq \alpha \leq 1, \alpha\rho_1 + (1-\alpha)\rho_2 \in \mathcal{S}(\mathcal{H}).$$

A state is mixed if $\mathrm{tr}\,\rho^2 < tr\rho$, and pure otherwise. Pure states are the extremal points of $\mathcal{S}(\mathcal{H})$. They are one-dimensional projectors $|\psi\rangle\langle\psi|$. It is common to identify the wave function $|\psi\rangle$ with the projector $|\psi\rangle\langle\psi|$ when talking of a pure state.

Parametrisation of states. For a finite-dimensional system, we have $\mathcal{H} = \mathbf{C}^{\otimes d}$ and $\mathcal{B}(\mathcal{H})$ is just the algebra of complex $d \times d$ matrices. $\mathcal{B}(\mathcal{H})$ is a vector space for the scalar product $\mathcal{B}(\mathcal{H}) \ni (A, B) \to \mathrm{tr}(A^*B) \in \mathbf{C}$. For $d = 2$, a useful parametrisation of $\mathcal{S}(\mathcal{H})$ is given by

$$\mathcal{S}(\mathcal{H}) = \{\rho = \frac{\mathbf{1}}{d} + \frac{1}{2}\mathbf{n}\cdot\sigma : \mathbf{n} \in \mathbf{R}^3, ||\mathbf{n}|| \leq 1\}, \tag{2.1}$$

where σ are the Pauli matrices:

$$\sigma = \sigma_x, \sigma_y, \sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, .$$

$\mathcal{S}(\mathcal{H})$ is then called the Bloch ball and its boundary, the set of pure states, is called the Bloch sphere.

Measurements and Measures

The mathematical description of a *measurement* in quantum mechanics is usually given by an *orthogonal* resolution of unity, i.e., a family of operators $\mathcal{M} = \{M_x :\ x \in X\}$ such that

1. the M_x's are orthogonal projectors, i.e., $M_x \geq 0,\ \ M_xM_{x'} = \delta_{xx'}M_x\ \ \forall x, x' \in X$;
2. The M_x's sum up to unity: $\sum_x M_x = \mathbf{1}_{\mathcal{H}}$,

where $\mathbf{1}_{\mathcal{H}}$ represents the identity operator on $\mathcal{H}$. $\mathcal{M}$ is sometimes called a von Neumann measurement. $\mathcal{M}$ is often called a von Neumann, or projection valued (pv), measure. This latter terminology comes from the fact that for a system prepared in a state ρ, the quantities $\mathrm{tr}(\rho M_x)$ define a probability measure over the set X:

$$0 \leq \mathrm{tr}(\rho M_x) \leq 1,\ \forall x \in X\ \text{ and }\ \sum_x \mathrm{tr}(\rho M_x) = 1, \tag{2.2}$$

in accordance with the axioms of quantum mechanics. However, a family of operators $\mathcal{O} = \{O_x :\ x \in x\}$ need not be a pv measure in order to satisfy condition (2.2). It is sufficient for $\mathcal{O}$ to satisfy

$$0 \leq O_x \leq \mathbf{1}_{\mathcal{H}}, \tag{2.3}$$

$$\sum_y O_x = \mathbf{1}_{\mathcal{H}}. \tag{2.4}$$

Such resolutions of unity are called positive operator valued (pov) *measures*.

Examples Consider a single qubit and let $\{\mathbf{n}_x : x \in X\}$ denote a set of unit vectors satisfying $\sum_x c_x \mathbf{n}_x = 0$, where the coefficients c_x satisfy $0 < c_x < 1 \ \forall x \in X$ and $\sum_x c_x = 2$. The family of operators

$$\mathcal{O} \equiv \{O_x = \frac{c_x}{2}(\mathbf{1} + \mathbf{n}_x \cdot \sigma)\}$$

defines a pov measure. Note that for $d = 2$, $Y = \{0, 1\}$ and $c_0 = c_1 = 1$, we just find a usual ("Stern-Gerlach") pv measure for a two-level system.

The reason for considering pov measures is that there are circumstances where they are more informative than pv measures. In the next chapter, we will talk of measurements allowing to guess optimally about a direction from 2 quantum systems (see Sect.3.4). In fact, the optimal is then achieved by a pov measure.

We have an idea of the *measurements* corresponding to a pv measure. These are characterised by observables of the form

$$A = \sum_x \lambda_x M_x,$$

where $\{\lambda_x\}$ are the real eigenvalues of A. But to what measurement does a povm correspond? The following theorem answers this question.

Theorem 2.2.1 (Neumark) *Let $\mathcal{O} = \{O_y : y \in Y\}$ denote a POVM over a Hilbert space $\mathcal{H}$. There exists a Hilbert space $\mathcal{H}'$ such that $\mathcal{O}$ can be realised by extending $\mathcal{H}$ to $\mathcal{H} \otimes \mathcal{H}'$, and performing a von Neumann measurement over $\mathcal{H} \otimes \mathcal{H}'$.*

2.3 Evolutions

We need a mathematical tool to describe quantum information processing. This tool is the concept of evolution, that is a map T taking density operators over a Hilbert space $\mathcal{H}$ into density operators over a Hilbert space $\mathcal{K}$. There are two (equivalent!) manners to describe an evolution.

Unitary operators

The evolution of the state of an isolated system is governed by a unitary operator. Indeed, let $|\psi(\mathbf{x}, t)\rangle$ denote the state of a quantum system at time t (the variable $\mathbf{x}$ represents the degrees of freedom of the system). From the Schrödinger equation

$$H|\psi\rangle = i\hbar\partial_t|\psi\rangle, \tag{2.5}$$

we see that

$$|\psi(\mathbf{x}, t)\rangle = e^{-i\int_{t_0}^{t} ds H(s)}|\psi(\mathbf{x}, t_0)\rangle \equiv U(t, t_0)|\psi(\mathbf{x}, t_0)\rangle. \tag{2.6}$$

Since H is self-adjoint, $U(t, t_0)$ is unitary. Therefore any evolution can be modelled by a unitary operator acting jointly on the system carrying the quantum information and an auxiliary system, often called the ancilla.

In quantum information, we will often only be interested in the "initial" state and in the "final" state of the evolution of a system. (Such a situation is common in classical information theory. When flipping a bit for example, we often don't need to know how the bit is actually flipped.) Therefore, when studying a quantum information process, we will often not need Schrödinger equation, and it will be enough to describe the evolution achieving this process in terms of a unitary operator, without caring too much about the Hamiltonian generating it.

CP maps

The second approach to describe the evolution of quantum systems, called "axiomatic" by some authors [18], consists in imposing a minimal set of requirements to make the map T, describing the evolution, consistent with the statistical interpretation of quantum mechanics. In particular, T must transform density operators into density operators. Hence, (i) T must respect convex combinations of states (ii) T must map positive operators to positive operators. In addition, the condition (ii) should also be satisfied if T is only applied to a part of a larger system. Finally, (iii) T has to respect normalisation. Let $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$ denote respectively the "input" and the "output" Hilbert space of the evolution T. Mathematically these conditions read:

1. T must be linear.
2. $T : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out})$ must be positive: $\forall A \in \mathcal{B}(\mathcal{H})$, if $A \geq 0$, then $T(A) \geq 0$. T must be completely positive: any extension of the form $T \otimes \mathbf{1}_{\mathcal{H}'} : \mathcal{B}(\mathcal{H}_{in}) \otimes \mathcal{B}(\mathcal{H}') \to \mathcal{B}(\mathcal{H}_{out}) \otimes \mathcal{B}(\mathcal{H}')$, where $\mathcal{H}'$ is an auxiliary Hilbert space, must be positive.
3. T must be trace-preserving: $\forall A \in \mathcal{B}(\mathcal{H}), \operatorname{tr} T(A) = \operatorname{tr} A$.

How can we reconcile the two apparently different descriptions of a quantum evolution? Actually, each one is adapted to a specific scheme. The cp maps formalism is well adapted to describe the evolution of an *open* system, whereas the unitary operators formalism describes the evolution of an *isolated* system. Fortunately, both approaches are equivalent. First, the unitary operator formalism is embedded in the cp map formalism. Indeed, let $\mathcal{H} = \mathcal{H}_{in} \otimes \mathcal{K} = \mathcal{H}_{out} \otimes \mathcal{K}'$ denote the Hilbert space on which the unitary operator responsible for the evolution acts. One can check that $\forall |\phi\rangle \in \mathcal{K}$, the mapping

$$T : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out}) : \rho \to \mathrm{Tr}_{\mathcal{K}'}(U\ (\rho \otimes |\phi\rangle\langle\phi|)\ U^*)$$

is indeed a trace-preserving cp map. Second, any cp map can be realised by applying a unitary operator acting jointly on the input system and an auxiliary system and by tracing over the ancilla. This statement is a consequence of the Stinespring dilation theorem for cp-maps [19, 20]. This theorem allows to represent the evolution of open systems, as unitary operators allow to represent the evolution of isolated systems. We state this theorem here for the convenience of the reader.

Theorem 2.3.1 (Stinespring) *Every completely positive map $T : \mathcal{B}(\mathcal{H}_{in}) \to \mathcal{B}(\mathcal{H}_{out})$ has the form*

$$T(X) = V^*(X \otimes \mathbf{1}_{\mathcal{K}})V, \tag{2.7}$$

with an additional Hilbert space $\mathcal{K}$ and an operator $V : \mathcal{H}_{out} \to \mathcal{H}_{in} \otimes \mathcal{K}$. This decomposition is unique up to unitary equivalence.

2.4 Bipartite systems and quantum entanglement

Quantum entanglement is the essential feature making quantum information theory distinct from classical information theory. It has long been considered only just as a consequence of the weird superposition principle of quantum mechanics. But now entanglement has been identified as the essential resource making possible applications such as quantum teleportation, efficient quantum computation, and to a certain extent quantum cryptography.

Let $\mathcal{H}_A \otimes \mathcal{H}_B$ denote the Hilbert space of a composite, spatially separated, quantum system. $\mathcal{H}_A$ is associated with the first quantum system, and $\mathcal{H}_B$ with the second. A bipartite density operator $\rho \in \mathcal{B}(\mathcal{H}_A) \otimes \mathcal{B}(\mathcal{H}_B)$ is called separable if it admits a convex decomposition into tensor products of density operators:

$$\rho = \sum_i p_i \, \rho_i^A \otimes \rho_i^B, \tag{2.8}$$

where the ρ_i^A's (resp. the ρ_i^B's) are density operators over $\mathcal{H}_A$ (resp. $\mathcal{H}_B$), and where the coefficients p_i form a probability distribution. An entangled state is a state which is not separable. Of course, these two definitions naturally generalise to n-partite systems. It is crucial to understand that an entangled state is not just a correlated state. There are separable states exhibiting correlations. For example, considering two qubits, the state $\frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \otimes |1\rangle\langle 1|$ shows correlations but is separable. The difference between separable and entangled states lies in the *nature* of correlations. *An entangled state cannot be prepared by means of local operations and classical communication.*

The definition (2.8) of separability and entanglement is conceptually sound but not very operational. Unfortunately, except in some specific situations, it is difficult to say much more: a central problem in quantum information theory is to provide operational measures to quantify entanglement. There are a few cases where this can be done easily, see for example [20] for a review. One case is particularly easy to consider; bipartite pure states, thanks to the following theorem.

Theorem 2.4.1 (Schmidt decomposition) *Let $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ denote a bipartite pure state (*$\dim \mathcal{H}_A$ *and* $\dim \mathcal{H}_A$ *need not match). There always exists an orthonormal basis $\{|a_i\rangle\}$ of $\mathcal{H}_A$ and an orthonormal basis $\{|b_i\rangle\}$ of $\mathcal{H}_B$, and positive coefficients $\{\lambda_i\}$ such that we can write*

$$|\Psi\rangle = \sum_i \sqrt{\lambda_i} |a_i\rangle |b_i\rangle \tag{2.9}$$

Proof. Let $\rho_A = \mathrm{Tr}_B \, |\Psi\rangle\langle\Psi|$ denote the reduced state of system A alone. We can always diagonalise ρ_A and write

$$\rho_A = \sum_i \lambda_i |a_i\rangle\langle a_i| \tag{2.10}$$

for some orthonormal basis $\{|a_i\rangle\}$ of $\mathcal{H}_A$. In full generality, we can thus write the expansion

$$|\Psi\rangle = \sum_i |a_i\rangle |\phi_i\rangle, \tag{2.11}$$

where the states $|\phi_i\rangle$ need not be orthogonal nor normalised. Thus, we have

$$\rho_A = \sum_{i,j} |a_i\rangle\langle a_j| \langle \phi_j | \phi_i \rangle. \tag{2.12}$$

Comparing (2.10) with (2.12), we see that $\langle \phi_j | \phi_i \rangle = \lambda_i \delta_{ij}$. Setting $|\phi_i\rangle = \sqrt{\lambda_i} |b_i\rangle$, we see that $\{|b_i\rangle\}$ is a set of orthonormal vectors of $\mathcal{H}_B$ and that Eq. (2.9) holds.□

It is easy to quantify entanglement with the Schmidt decomposition. A crude measure is provided by the Schmidt number [21]: it is just the number of non-zero coefficients in the Schmidt decomposition of $|\Psi\rangle$. A more subtle measure is given by the von Neumann entropy of either part of $|\Psi\rangle$:

$$S(\rho_A) \equiv -\operatorname{tr} \rho_A \log \rho_A = S(\rho_B) = -\operatorname{tr} \rho_B \log \rho_B. \tag{2.13}$$

Eq. (2.9) makes it easy to compute $S(\rho_A)$. We have

$$S(\rho_A) = H(\{\lambda_i\}) = \sum_i \lambda_i \log \lambda_i. \tag{2.14}$$

The interpretation of $S(\rho_A)$ as a measure of entanglement is then straightforwardly derived from the interpretation of Shannon entropy.

Consider two extreme situations. For a separable state, $|\phi_A\rangle|\phi_B\rangle$, there exists a basis of $\mathcal{H}_A$ (resp. $\mathcal{H}_B$) containing $|\phi_A\rangle$ (resp. $|\phi_B\rangle$), such that the outcome of a measurement, whose associated observable is of the form $x_A|\phi_A\rangle\langle\phi_A| + \text{Rest}$ (resp. $x_B|\phi_B\rangle\langle\phi_B| + \text{Rest}$), can be predicted with certainty. For a maximally entangled state $\frac{1}{\sqrt{d}}\sum_i |a_i\rangle|b_i\rangle$, where $d = min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$, the outcome is random whatever the basis, and has entropy $\log d$. This entropy is maximal for the smaller Hilbert space. This basic example is sufficient to show what is so special about entanglement. There is certainly some information encoded in a bipartite state $|\Psi\rangle$: there certainly exists a basis of $\mathcal{H}_A \otimes \mathcal{H}_B$ containing $|\Psi\rangle$. However, for an entangled state, this information cannot be fully extracted *locally*. In particular, for a maximally entangled state, no information can be extracted locally: all the information content of $|\Psi\rangle$ lies in non-local correlations.

Example. The simplest example of an entangled state is given by the Bell (or EPR) pure state

$$|\phi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}). \tag{2.15}$$

We have $\text{Tr}_B\, |\phi_0\rangle\langle\phi_0| = \text{Tr}_A\, |\phi_0\rangle\langle\phi_0| = \frac{1}{2}$ and $S(\rho_A) = 1$ bit. We will see in chapter 6 that this state plays a crucial role in quantum cryptography. The reason is that if two parties share a system in the state $|\phi_0\rangle$, they can securely extract a common secret bit from it.

2.5 Quantum error correction

Quantum error correcting codes are sophisticated procedures allowing the protection of quantum information against environmental noise. Without them, there would be no hope of building a quantum computer someday. The main reason for us to study quantum error correcting codes is that they provide a very powerful formalism to assess the security of quantum cryptographic protocols, which will be extensively used in chapter 6.

Necessary and sufficient condition for quantum error correction

Let us first consider the case of one qubit (with Hilbert space $\mathcal{H}$), subject to interactions with its environment (with Hilbert space $\mathcal{H}_E$). Without loss of generality, we can suppose that the environment is initially in some pure state $|0\rangle_E$. The evolution of this qubit and its environment can be described in terms of (some lines of) a unitary operator:

$$\begin{aligned} U : \mathcal{H} \otimes \mathcal{H}_E \to \mathcal{H} \otimes \mathcal{H}_E : \quad & |0\rangle|0\rangle_E \quad \to |0\rangle|e_{00}\rangle_E + |1\rangle|e_{01}\rangle_E \\ & |1\rangle|0\rangle_E \quad \to |0\rangle|e_{10}\rangle_E + |1\rangle|e_{11}\rangle_E. \end{aligned} \tag{2.16}$$

$$\tag{2.17}$$

Note that the four states $|e_{ij}\rangle$ need not be normalised or mutually orthogonal. By linearity, the evolution of a state $|\psi\rangle = a|0\rangle + b|1\rangle$ is given by

$$|\psi\rangle|0\rangle_E \to \mathbf{1}|\psi\rangle|e_{\mathbf{1}}\rangle + X|\psi\rangle|e_X\rangle + Y|\psi\rangle|e_Y\rangle + Z|\psi\rangle|e_Z\rangle, \tag{2.18}$$

where X, Y, Z is a shorthand notation for the three Pauli matrices and where the states $\{|e_{\mathbf{1}}\rangle, |e_X\rangle, |e_Y\rangle, |e_Z\rangle\}$ are linear combinations of the states $\{|e_{ij}\rangle\}$ (independent of $|\psi\rangle$).

Eq (2.18) allows us to "interpret" the effect of the interaction of the qubit with its environment. The qubit might be left unaffected ($\mathbf{1}$), undergo a bit-flip $a|0\rangle + b|1\rangle \rightarrow b|0\rangle + a|1\rangle$ (X), a phase-flip $a|0\rangle + b|1\rangle \rightarrow a|0\rangle - b|1\rangle$ (Z), or both $a|0\rangle + b|1\rangle \rightarrow b|0\rangle - a|1\rangle$ ($Y = iXZ$)[1]. Rigorously, such an interpretation doesn't hold because the states $\{|e_{\mathbf{1}}\rangle, |e_X\rangle, |e_Y\rangle, |e_Z\rangle\}$ are in general not mutually orthogonal. Still, we can keep the important property that the unitary U can be decomposed in terms of the Pauli operators $\{\mathbf{1}, X, Y, Z\}$.

Similarly, the interaction of an n-qubit system in a pure state $|\psi\rangle$ can be modelled as

$$|\psi\rangle|0\rangle_E \rightarrow \sum_a E_a|\psi\rangle|e_a\rangle_E, \tag{2.19}$$

where the operators E_a are tensor products of Pauli operators, i.e.

$$G_n \equiv \{\mathbf{1}, X, Y, Z\}^{\otimes n}. \tag{2.20}$$

Eq (2.19) constitutes our error model. The set of *correctable* errors is a subset $\mathcal{E} \subseteq G_n$. Each Pauli operator E_a can be assigned a weight, that is an integer $0 \leq t \leq n$. This weight is the number of qubits on which E_a acts non-trivially. Again, one can "interpret" the weight of a Pauli operator as the number of qubits on which an error occurs. Typically, in quantum error correction, one takes $\mathcal{E}$ to be the set of Pauli operators of weight up to t.

A quantum code, $\mathcal{Q} \subseteq \mathcal{H}^{\otimes n}$, is a subspace of a Hilbert space, the "logical qubits subspace". A quantum code is robust against errors in $\mathcal{E}$ if there exists an auxiliary system A (an ancilla), and a recovery operator $R \in \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}_A)$ such that

$$R \otimes \mathbf{1}_E : \mathcal{H} \otimes \mathcal{H}_E \otimes \mathcal{H}_A \rightarrow \mathcal{H} \otimes \mathcal{H}_E \otimes \mathcal{H}_A : \sum_a E_a|\psi\rangle \otimes |e_a\rangle|\phi\rangle_A \rightarrow$$

$$\sum_{\mu,a} (R_\mu \otimes \mathbf{1}_E)(E_a \otimes \mathbf{1}_A)|\psi\rangle|e_a\rangle_E|\mu\rangle_A \rightarrow |\psi\rangle \otimes |whatever\rangle_{EA}, \tag{2.21}$$

where $|\phi\rangle$ is some initial state of the ancilla and where $\{|\mu\rangle\}$ is an orthonormal basis of $\mathcal{H}_A$. We now give a necessary and sufficient condition for error recovery to be possible.

Theorem 2.5.1 (Condition for quantum error correction) *[22] Let $\mathcal{H}^{\otimes n}$ denote the Hilbert space of n qubits. Let $\mathcal{Q} \subseteq \mathcal{H}^{\otimes n}$ denote a quantum code, and let $\{|\bar{i}\rangle\}$ denote an orthonormal basis of $\mathcal{Q}$. $\mathcal{E}$ is a set of correctable errors if and only if*

$$\langle \bar{j}|E_b^* E_a|\bar{i}\rangle = C_{ba}\delta_{ij}, \quad \forall E_a, E_b \in \mathcal{E},\ \forall i, j. \tag{2.22}$$

The fact that this conditions is necessary is evident. Would we have $\langle \bar{j}|E_b^* E_a|\bar{i}\rangle \neq 0$ for some $E_b, E_a \in \mathcal{E}$ and some $i \neq j$, the errors would destroy the distinguishability between orthogonal codewords[2], and quantum information would certainly be damaged. It is also easy to understand that

$$\langle \bar{j}|E_b^* E_a|\bar{i}\rangle = \delta_{ba}\delta_{ij}, \quad \forall E_a, E_b \in \mathcal{E},\ \forall i, j \tag{2.23}$$

[1] the i phase can be absorbed in the definition of $|e_Y\rangle_E$.

[2] The codewords of a quantum code are just the vectors of this code.

is a sufficient condition. Eq (2.23) states that distinct errors should be distinguishable. Now the fact that Eq. (2.22) is already a sufficient condition for recovery is a bit counter-intuitive but quite straightforward to prove ([22, 21]). A quantum code that satisfies the condition (2.22) but not the condition (2.23) is said to be degenerate.

A *binary* quantum code is a 2^k-dimensional logical subspace embedded into a 2^n-dimensional Hilbert space and can be represented in terms of qubits. The *distance* d of a binary quantum code is the minimum weight of a Pauli operator E such that

$$\langle \bar{i}|E|\bar{j}\rangle \neq C\delta_{ij}. \tag{2.24}$$

Similarly to classical error correction, it is usual to refer to a binary quantum code by a triple $[[n, k, d]]$. (The double bracket notation is only used to distinguish the notation of quantum codes from that of classical codes.)

Mimicking again the classical theory, we say that a binary quantum code corrects t errors if the set $\mathcal{E}$ of correctable errors includes all Pauli operators of weight t or less. In virtue of our definition of the distance of a code, a quantum error correcting code with distance $d = 2t + 1$ can correct t errors. Also, a quantum code with distance $d = t + 1$ can correct t errors at known locations.

Stabiliser Codes

The main idea of stabiliser codes can be understood very intuitively with the following scheme. Take a Hilbert space $\mathcal{K}$, typically $\mathcal{K}$ will be the Hilbert space of n qubits, which has the direct sum structure:

$$\mathcal{K} = \mathcal{H}_0 \oplus \ldots \oplus \mathcal{H}_{m-1}, \tag{2.25}$$

where all the Hilbert spaces $\mathcal{H}_i$ are isomorphic ($\dim \mathcal{H}$ divides $\dim \mathcal{K}$). $\mathcal{H}_0$ is a quantum code. This code is robust against the errors that map $\mathcal{H}_0$ onto one of its copies but "doesn't affect its internal structure". That is, if $\{|v_{kl}\rangle\}$ denotes an orthonormal basis of $\mathcal{H}_k$ ($l = 0 \ldots \dim_{\mathcal{H}_k}$), the errors of the form

$$E : \mathcal{H}_k \to \mathcal{H}_{k'} : |\psi_k\rangle = \sum_l c_l |v_{kl}\rangle \to |\psi'_k\rangle = \sum_l c_l |v_{k'l}\rangle \tag{2.26}$$

can be diagnosed and reversed without damaging the quantum information contained in $|\psi_k\rangle$. The whole problem of quantum error correction is to find clever decompositions of the form Eq. (2.25), so that the set of correctable errors is physically sensible.

Let us now be more precise. Let $S = \{F_\alpha\}$ denote an Abelian subgroup of G_n, which we call the *stabiliser*. A *stabiliser code*, $\mathcal{H}_S \subset \mathcal{H}^{\otimes n}$, is a simultaneous eigenspace of all operators $\{F_\alpha\}$ with a fixed set of eigenvalues $\{\lambda_i\}$. To make the presentation simple, let us assume that $\mathcal{H}_S$ is such that $\lambda_i = 1\ \forall i$. Other equivalent stabiliser codes, associated with a different set of eigenvalues, can be defined by "translation" from $\mathcal{H}_S$. The stabiliser thus preserves all codewords:

$$F_\alpha|\psi\rangle = |\psi\rangle\ \forall |\psi\rangle \in \mathcal{H}_S,\ \forall F_\alpha \in S. \tag{2.27}$$

A set of generators for S, $\mathfrak{g} = \{M_i\} \subset S$, is a set of independent operators (none of them can be expressed as a tensor product of the others) and such that each element of S can be expressed as a product of generators. One can prove that if S has $n - k$ generators, then $\dim \mathcal{H}_S = 2^k$. $\mathfrak{g}$ is the set of "check operators" for the code, the collective observables that we measure to diagnose the errors. If $M_i = 1$ for all $M_i \in \mathfrak{g}$, then we judge that the information is undamaged. If there is an operator M_i such that $M_i = -1$, then an error is detected (all elements of G_n have eigenvalue either 1

or -1). Recall that each error operator can be expanded in terms of a tensor product of Pauli operators. Hence, we have

$$M_i E_a = (-)^{s_i^a} E_a M_i \tag{2.28}$$

The s_i^a's, $i = 1 \dots n-k$ constitute the syndrome of the error E_a. Now, given a stabiliser code, we have the following sufficient condition for the error correction condition (2.22) to hold:

> $\forall E_a, E_b \in \mathcal{E}$, either $E_a^* E_b \in S$, either $\exists M_i \in \mathfrak{g}$ such that M_i and $E_a^* E_b$ anticommute.

One way to prove this assertion is to show that it provides us with an algorithmic means to operate error correction. Let a codeword $|\psi\rangle$ undergo an error

$$\mathcal{E} \ni E_b : |\psi\rangle \rightarrow E_b|\psi\rangle.$$

Apply

$$\mathcal{E} \ni E_{a_1}^* : E_b|\psi\rangle \rightarrow E_{a_1}^* E_b|\psi\rangle$$

for some $E_{a_1}^*$. Then measure all stabiliser generators $M_i \in \mathfrak{g}$. For all M_i, $E_{a_1}^* E_b$ and M_i either commute or anticommute[3]. If one of the M_i is such that $M_i \neq 1$, then apply

$$\mathcal{E} \ni E_{a_1} : E_{a_1}^* E_b|\psi\rangle \rightarrow E_b|\psi\rangle.$$

Repeat the procedure with $E_{a_2}, \dots$ until one encounters E_{a_j} such that $[E_{a_j}^* E_b, M_i] = 0 \; \forall M_i \in \mathfrak{g}$, that is, $E_a^* E_b \in S$, so that the damaged state $E_b|\psi\rangle$ is transformed to the undamaged state $E_{a_j}^* E_b|\psi\rangle = |\psi\rangle$.

Example: Shor's 9 qubit code. The simplest example of a quantum code is Shor's $[[9, 1, 3]]$ stabiliser code [21]. This code is a natural quantum adaptation of the 3-bit majority voting classical code. Indeed the codewords of this code are the linear span of the two logical qubit states

$$|0'\rangle = \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \tag{2.29}$$

$$|1'\rangle = \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \tag{2.30}$$

Denoting $X_i \equiv \mathbf{1}^{\otimes i-1} \otimes X \otimes \mathbf{1}^{n-i}$ and $Z_i \equiv \mathbf{1}^{\otimes i-1} \otimes Z \otimes \mathbf{1}^{n-i}$, eight stabiliser generators of this code are for example:

$$Z_1 Z_2, Z_2 Z_3, Z_4 Z_5, Z_5 Z_6, Z_7 Z_8, Z_8 Z_9, \tag{2.31}$$

which allow to diagnose a single bit flip error, and the operators

$$X_1 X_2 X_3 X_4 X_5 X_6, X_4 X_5 X_6 X_7 X_8 X_9, \tag{2.32}$$

which allow to diagnose in which cluster of 3 qubits a single phase flip might have occured. Note that this code is degenerate. For example, if a phase flip error has occured on an arbitrary qubit of the first cluster of a state $a|0'\rangle + b|1'\rangle$, this state will be transformed to

$$\frac{1}{2^{3/2}}(a(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$$

[3]So, up to a phase, measuring the stabiliser generators doesn't affect the state$E_{a_1}^* E_b|\psi\rangle$

$$+b(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)), \tag{2.33}$$

whatever the precise error Z_1, Z_2 or Z_3 occured. This error will be diagnosed with the operators (2.32) and corrected upon applying $X_1X_2X_3$. This is consistent with the condition for error correction since $(X_1X_2X_3)Z_1$, $(X_1X_2X_3)Z_2$, $(X_1X_2X_3)Z_3$ all belong to the stabiliser.

2.6 Summary

In summary, we have reviewed some fundamental concepts and results of quantum information theory: povm's, a model of evolution for quantum information processing, quantum entanglement, and quantum codes. An appendix on classical information theory is also presented below.

There are three excellent references for further reading. The first one are the notes of Preskill's course on quantum information and quantum computation [21], the second are Keyl's notes [20], and the third is the book of Alber et al. [3]. There, one can find a more detailed description of the matters discussed in this chapter. Moreover, important issues, left aside here, are discussed. An account on quantum computation can be found in [21] or [3], for example, and [20] provides a quite detailed review of entanglement measures and coding theorems for quantum channels. Ref.[3] also deals with experimental aspects of quantum information.

2.7 Appendix: Classical Information Theory

The foundation of classical information theory is the study of two problems: How can we efficiently compress a message? This problem is called "source coding". And how can we reliably communicate over a noisy channel? This problem is called "channel coding".

Source coding

Shannon entropy. A message is a string of "symbols" drawn from an "alphabet" $\{x_0 \dots x_{m-1}\}$. We suppose that the symbols in the message are statistically independent[4], and that each x_k occurs with a probability p_k. We call ensemble an alphabet together with the probabilities associated to each of its symbols: $\{x_k, p_k\}$.

Let us consider the simplest example of a binary alphabet, where the symbol '0' occurs with a probability $(1-p)$ and consider messages of size n. In virtue of the law of large numbers, when $n >> 1$, *typical* messages will contain (about) $n(1-p)$ 0's and (about) np 1's. The key idea of compression is to ignore other messages, since their probability of occurrence decays exponentially fast with n. We then have a reduction from the set $\{0,1\}^n$ containing 2^n strings to the set $\mathcal{T}_n(p)$ of typical messages. The gain in compression is evaluated from calculating the size of $\mathcal{T}_n(p)$. The number of typical messages is given by the binomial coefficient $\binom{n}{np}$. Using Stirling approximation $\log n! = n\log n - n + O(\log n)$, we get

$$\log \binom{n}{np} \approx n\log n - n - (np\log np - np + n(1-p)\log n(1-p) - n(1-p)) = nh(p),$$

[4]We make this hypothesis to simplify the analysis, but it sometimes has to be relaxed. In the English language for example, there are certainly correlations between the symbols of a message: the probability to have four consecutive consonants is for example very low.

where

$$h(p) = -p \log p - (1-p) \log(1-p) \tag{2.34}$$

is the binary Shannon entropy. Shannon entropy is the fundamental tool of information theory. It tells us how many bits of information are, on average, carried by each bit of a message. Stated otherwise, $h(p)$ quantifies our a priori ignorance about the outcome of the statistical experiment:"draw randomly a symbol from the alphabet".

The reasoning generalises straightforwardly to an m-symbol alphabet. The number of typical strings is now given by

$$\frac{n!}{\pi_k (np_k)!} \approx 2^{nH(X)}, \tag{2.35}$$

where $H(X) = \sum_k -p_k \log p_k$ is the Shannon entropy of the ensemble $X = \{x_k, p_k\}$. Let us now state the first fundamental result discovered by Shannon.

Theorem 2.7.1 (Shannon's source coding theorem) *[1] Let $X = \{x_k, p_k\}$ denote an ensemble. (i) There exists a sequence of codes Γ_μ, compressing n_μ-bit strings into k_μ-bit strings, $\lim n_\mu \to \infty = \infty$, such that compression is asymptotically effected without loss of information, and such that the rate of compression, $\lim_{\mu\to\infty} k_\mu/n_\mu$ is asymptotically given by the Shannon entropy $H(X)$. (ii) $H(X)$ is the optimal rate of lossless compression of the source X.*

Channel coding

Mutual information. The mutual information $I(X : X')$ quantifies the correlations between the messages drawn from two ensembles X^n and X'^n. Consider a situation where an emitter wants to transmit a message $\mathbf{m} = m_0 \dots m_{n-1}$ to a receiver. If the channel connecting them is noisy, the receiver will in general get a different message $\mathbf{m}' = m'_0 \dots m'_{n-1}$. $I(X : X')$ quantifies the information about $\mathbf{m}$ which is gained from the knowledge of $\mathbf{m}'$.

The noisy channel can be characterised by the conditional probabilities $p(x'|x)$, i.e. the probabilities that $x' \in \{x'_k\}$ is received when $x \in \{x_k\}$ was sent. Before receiving x', the symbol sent by the emitter is, from the point of view of the receiver , characterised by the entropy $h(p)$. Now the knowledge of x' allows for a re-evaluation of the probability distribution for x. Indeed, we have (Bayes' theorem):

$$p(x|x') = \frac{p(x'|x)p(x)}{p(x')},$$

so that *after* having received the symbol x', the message sent by the emitter is, from the point of view of the receiver, characterised by the entropy

$$\begin{aligned} H(X|X') = \sum_{k'} p_{k'} H(X|X' = x'_k) &= -\sum_{k'} p_{k'} \sum_k p(k|k') \log p(k|k') \\ &= H(X, X') - H(X), \end{aligned} \tag{2.36}$$

where $p(k|k')$ represents the probability that the source X has emitted the symbol x_k when the source X' has emitted the symbol x'_k, and where $H(X, X')$ represents the Shannon entropy of the joint probability distribution of (X, X'). $H(X|X')$ can be interpreted as the residual ignorance about X when X' is known. From $H(X)$ and $H(X|X')$, we define the mutual information

$$I(X : X') \equiv H(X) - H(X|X') \tag{2.37}$$

$$= H(X') - H(X'|X). \tag{2.38}$$

$I(X : X')$ is the reduction of entropy about X brought by the knowledge of X'. Re-expressing

$$I(X : X') = \sum_{k,k'} p_{kk'} \log \frac{p_{kk'}}{p_k p_{k'}},$$

we see that the mutual information quantifies the "distance" between the two probability distributions $p_{kk'}$ and $p_k p'_k$. Clearly, $I(X : X')$ equals 0 if the variables X and X' are not correlated.

Given a noisy channel exhibiting correlations between the input and the output, it is always possible to use it for reliable communication with an appropriate *code*, whose role is to introduce enough redundancy to ensure the reliability of the communication. The simplest code for reliable communication to think of is the majority vote coding: the emitter transmits 0^n when she wants to communicate the symbol '0' and 1^n when she wants to communicate the symbol '1'. But does every noisy channel allow reliable communication? At least asymptotically. That is, can we, for every noisy channel, devise a sequence of codes C_α, transmitting k_α bits of communication with n_α encoding bits, $\lim_{\alpha\to\infty} n_\alpha = \infty$, so that the limit of the rates of communication of this sequence, $\lim_{\alpha\to\infty} k_\alpha/n_\alpha$ tends to a non-zero value? The answer to this question is the content of Shannon's second theorem.

Theorem 2.7.2 (Shannon's channel coding theorem) *[1] Let T denote a channel, whose input is represented by an ensemble X and output by an ensemble X'. The capacity of T, that is the optimal rate of communication over T defined as*

$$C(T) = sup \lim_{n\to\infty} \frac{k}{n}, \tag{2.39}$$

where the supremum is taken over all sequences of codes, is given by

$$C(T) = sup_{\{p(x_k)\}} I(X : Y). \tag{2.40}$$

where $\{p(x_k)\}$ is the probability distribution associated to the ensemble X.

Classical linear codes

Shannon's second theorem informs us on what is the maximal rate of transmission of a channel effecting stochastic errors, but it doesn't give any hint on how to construct a code ensuring reliable communication. We conclude this survey of classical information theory with a description of an important class of error correcting codes: linear codes.

Let $\mathbf{F}_2^n$ denote the vector space of n-component bit vectors. The addition of two vectors in this space is defined component-wise and modulo 2. Furthermore, we endow this vector space with the scalar product $\mathbf{v} \cdot \mathbf{w} = (\sum_i v_i w_i)$ mod2. A classical linear code C is a k-dimensional subspace of $\mathbf{F}_2^n$. The vectors of C are called the codewords. The weight of a binary string is simply the number of components of this string different from '0'. The distance of a code C is the minimum of the weights of the codewords of C.

The code orthogonal to C is defined by $C^\perp = \{w \in \mathbf{F}_2^n | w \cdot v = 0 \ \ \forall v \in C\}$. Note that $C \cap C^\perp$ may not be trivial. A parity check matrix for C is an $(n-k) \times n$ matrix H whose rows are a basis of $C^\perp$. An l-bit string v is a codeword of C iff

$$Hv = 0. \tag{2.41}$$

If a codeword v encounters an error $v \to v + \epsilon$, then the condition (2.41) is no longer satisfied[5]. The bit string $H\epsilon$ is called the syndrome of the error ϵ.

[5]Unless the error pattern ϵ is itself a codeword, in which case the error remains undetected.

Let $\mathcal{E}$ denote the set of errors we wish to be able to correct. Error recovery is possible if distinct errors give distinct syndromes. A linear code with distance $d = 2t+1$ can correct t errors occurring on arbitrary bits. Indeed, if ϵ_1 and $\epsilon_2 \neq \epsilon_1$ are two errors of weight at most t, $\epsilon_1 + \epsilon_2$ has a weight at most $2t$ and thus cannot be a codeword of C. Hence $H(\epsilon_1 + \epsilon_2) \neq 0$, that is $H\epsilon_1 \neq H\epsilon_2$: to different errors correspond different syndromes. Similarly, one proves that a linear code with distance $d = t+1$ can correct t errors on located bits. It is usual to refer to a classical code by a triple $[n, k, d]$ where n denotes the dimension of the space embedding the code, k the dimension of the code, and d its distance.

Finally, we note that the condition $Hw = s$ defines $\forall s \in \mathbf{F}_2^{n-k}$ a code C_s which is equivalent to C. We say that C_s is obtained by 'translation' from the code C: $C_s = \{w \in \mathbf{F}_2^n | Hw = s\}$.

Example 1: the three-bit code. The simplest code to think of is the 3-bit "majority voting" code $[3, 1, 3]$: $C = \{000, 111\} \subset \mathbf{F}_2^3$. A parity check matrix of this code is

$$H = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Suppose, for example, that the codeword 000 has encountered the error $000 \rightarrow 001$. H allows to diagnose that there is no error in the first block of two bits, but one error in the last block, i.e. the third bit has been affected. We can thus undo the error: $001 \rightarrow 000$.

Example 2: The Hamming code. The Hamming code is defined by the following parity check matrix:

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{2.42}$$

The distance of this code is d=3. First note that the vector $v = (1110000)$ passes the binary check matrix test $Hv = 0$. No codeword is of weight 1 because all columns of H are non-zero (if w_1 is of weight-1, then Hw_1 is a column of H). Neither is there any weight-2 codeword because because any weight-2 vector w_2 can be decomposed as a sum of two weight-one vectors, $w_2 = w_1 + w_1'$. Hence, $Hw_2 = Hw_1 + Hw_1' \neq 0$ because all columns of H are distinct. Thus the Hamming code is a [7,4,3] code: it encodes four bits into seven, and it is resistant to any single error ($2t + 1 = 3$).

The rows of H pass the parity check. Thus, if G denotes a generator for the Hamming code, we have $H \subset G$. Since $v = (1110000)$ lies in the code, and cannot be decomposed as linear combinaison of the rows of H , a possible generator matrix is

$$G = \begin{pmatrix} H \\ (v) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.43}$$

The dual of the Hamming code is the $[7, 3, 4]$ code generated by H. Thus we are in a situation where $C^\perp \subset C$, where C and $C^\perp$ denote respectively the Hamming code and its dual. In fact, the Hamming code is the union of two the cosets of $C^\perp$ in C:

$$C = \{C^\perp\} \cup \{v + C^\perp\}. \tag{2.44}$$

Chapter 3

Quantum cloning of finite dimensional systems

We first review the no-cloning theorem, whose importance in quantum information theory is central. We then discuss approximate quantum cloning machines. Afterwards, we study a special class of quantum cloning machines taking orthogonal qubits as input. We show that they can outperform standard quantum cloning machines, and discuss their implementation.

3.1 The no-cloning theorem

Let $\mathcal{H}$ denote the Hilbert space of some quantum system and let $\mathcal{H} \supset \mathcal{S} = \{|\psi_i\rangle : i \in \mathcal{I}\}$ denote a set of pure states. The set $\mathcal{I}$ of indices may be discrete or continuous. A quantum cloning machine is a device that takes a quantum system in some unknown state $|\phi\rangle \in \mathcal{S}$ and outputs two quantum systems, the clones. Ideally, each output should be in the state $|\phi\rangle$ too. Let $\rho_1(\phi)$ (resp. $\rho_2(\phi)$) denote the (possibly mixed) state of the first (resp. second) clone. If the quality of each clone is measured by the fidelities

$$f_1 = \langle\phi|\rho_1(\phi)|\phi\rangle,$$

$$f_2 = \langle\phi|\rho_2(\phi)|\phi\rangle,$$

that is by the overlap between the clones and the original, a perfect quantum cloning machine would make $f_1(\phi) = f_2(\phi) = 1$, $\forall|\phi\rangle \in \mathcal{S}$, leading to $\rho_1(\phi) = \rho_2(\phi) = |\phi\rangle\langle\phi|$. As the following theorem shows, this is not possible if the set $\mathcal{S}$ contains non-orthogonal states [23, 24].

Theorem 3.1.1 (No-cloning) *Let $\mathcal{H}$ and $\mathcal{S}$ be defined as above. Let also $\mathcal{K} = \mathcal{H}\otimes\mathcal{K}'$ denote the Hilbert space of some auxiliary system, and let $|\kappa\rangle \in \mathcal{K}$ denote some fiducial state.*

Either the set $\mathcal{S}$ is such that

$$|\langle\psi_i|\psi_j\rangle| = \delta_{ij} \quad \forall i, j \in \mathcal{I},$$

or there are no auxiliary Hilbert spaces $\mathcal{K}, \mathcal{K}'$ and auxiliary state $|\kappa\rangle \in \mathcal{K}$ such that the transformation

$$\mathcal{QC} : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H}^{\otimes 2} \otimes \mathcal{K}' : |\psi_i\rangle \otimes |\kappa\rangle \to |\psi_i\rangle^{\otimes 2} \otimes |\kappa_i'\rangle \tag{3.1}$$

can be performed exactly for all $i \in \mathcal{I}$.

Proof. Since $\mathcal{QC}$ should be achieved by a unitary transformation, it would have to satisfy

$$|\langle\psi_i|\psi_j\rangle| = |\langle\psi_i|\psi_j\rangle|^2|\langle\kappa_i'|\kappa_j'\rangle|. \tag{3.2}$$

If $|\langle\psi_i|\psi_j\rangle| = \delta_{ij}$ for all pair $|\psi_i\rangle, |\psi_j\rangle \in \mathcal{S}$, this condition can be satisfied. Now, suppose that $\mathcal{QC}$ can also be performed even if $\mathcal{S}$ contains one pair of non-orthogonal states $|\psi_k\rangle, |\psi_l\rangle$. We have for this pair:

$$0 < |\langle\psi_k|\psi_l\rangle| < 1. \tag{3.3}$$

On the other hand, Eq.(3.2) and the Schwarz inequality ($|\langle\kappa_i'|\kappa_j'\rangle|^2 \leq ||\kappa_i'||\ ||\kappa_j'|| = 1$) imply that

$$|\langle\psi_k|\psi_l\rangle| \geq 1, \tag{3.4}$$

leading to a contradiction. □

The no-cloning theorem marks a fundamental difference between classical and quantum information. However, its content shouldn't sound too astonishing. After all, if it were possible to perfectly duplicate the unknown state of a quantum system, it would be possible to perform measurements on the copy, and acquire information about the state of the original system without perturbing it, in contradiction with Quantum Mechanics.

The impossibility to perfectly clone non-orthogonal states puts strong limitations on the way we can encode quantum information, as is shown by the *no-cloning bound* in quantum error correction. In the classical case, we know that $[3, 1, 3]$ linear codes encoding one bit in three bits, robust against any error on a single bit, exist. The "majority vote" encoding $'0' \rightarrow' 000',' 1' \rightarrow' 111'$ does the job, and three is the minimum number of encoding bits of a classical code necessary to protect against a single arbitrary error. One may ask a similar question for quantum codes: How many encoding qubits are necessary to protect an encoded qubit against an arbitrary error occurring on a single encoding qubit? The no-cloning theorem tells us that we need at least five qubits [21]. Indeed, suppose that four qubits were sufficient: suppose that a $[[4, 1, 3]]$ quantum code exists. We could use this code to correct two errors at known locations, because we know that a code correcting t errors at any sites, also works to correct $2t$ errors at known sites. But then the following procedure would allow to achieve perfect quantum cloning: (i) Encode the state $|\psi\rangle$ to clone with the $[[4, 1, 3]]$ code:

$$E : \mathcal{H} \ni |\psi\rangle \rightarrow |E(\psi)\rangle \in \mathcal{H}^{\otimes 4}.$$

(ii) Split $|E(\psi)\rangle$ into two 2-qubit subsystems, say A and B. (iii) To the subsystem A (resp. B), append another 2-qubit subsystem Z_A (resp. Z_B) in the state $|00\rangle$. The state of the system AZ_A (resp. BZ_B) can now be seen as the state $|E(\psi)\rangle$ having undergone damage at two known sites. This damage could thus be reversed by the $[[4, 1, 3]]$ code: (iv) $|\psi\rangle \rightarrow |E(\psi)\rangle|E(\psi)\rangle$ (v) Decode the states $|E(\psi)\rangle_{AZ_A}$ and $|E(\psi)\rangle_{BZ_B}$ yielding two perfect clones, which is impossible. Hence no $[[4, 1, 3]]$ code exists. Note that five is an achievable bound for the number of qubits necessary for encoding one qubit with a code protecting against an arbitrary single qubit damage since we have explicit constructions for (perfect) $[[5, 1, 3]]$ quantum codes [25].

Remarkably, the no-cloning theorem can also be turned into a resource, as is beautifully demonstrated by quantum cryptography. Two authorised parties, willing to communicate privately, can exploit the impossibility to perfectly clone non-orthogonal states to encode classical information in quantum systems. Any intervention of a potential eavesdropper will cause a disturbance, that is, the eavesdropper will be detected (see chapters 5 and 6).

In the case where $\mathcal{S}$ is made of mutually orthogonal states, it is easy to achieve quantum cloning because there exists a measurement that discriminates amongst the states of $\mathcal{S}$. Any observable of the form $\sum_{i\in\mathcal{I}} \lambda_i |\psi_i\rangle\langle\psi_i|$ will work (eigenvalues are irrelevant, we just ask for non-degeneracy: $i \neq j \Rightarrow \lambda_i \neq \lambda_j$). The cloning procedure is then very simple: estimate the state to clone with an appropriate measurement and prepare two copies of it.

3.2 The universal duplicator of qubits

After the no-cloning theorem, it might seem useless to further study quantum cloning. The issue of approximate quantum cloning is however crucial. *Approximate* quantum cloning machines informs us on how the information contained in a quantum system can be distributed into other quantum systems.

Let us consider the case where the state to copy can be any pure qubit state belonging to $\mathcal{H} = \mathbf{C}^2$. Let $\{|0\rangle, |1\rangle\}$ denote as usual an orthonormal basis of $\mathcal{H}$. The simplest cloning transformation or "machine" one can think of is the Wootters-Zurek machine (WZ) [23], which acts perfectly on $|0\rangle$ and $|1\rangle$. This machine is defined by:

$$|0\rangle_A|Q\rangle_Z \rightarrow |\Phi(0)\rangle_{ABX} = |0\rangle_A|0\rangle_B|Q_0\rangle_X,$$

$$|1\rangle_A|Q\rangle_Z \rightarrow |\Phi(1)\rangle_{ABX} = |1\rangle_A|1\rangle_B|Q_1\rangle_X,$$

where A refers to the original state to clone, and Z some auxiliary system including the copy system, B, and an ancilla X. $|Q\rangle$ denotes some normalised fiducial state of the system Z, and $|Q_0\rangle$ (resp. $|Q_1\rangle$) denotes the normalised output state of the ancilla X when the state $|0\rangle$ (resp. $|1\rangle$) is copied.

Let $|\phi\rangle = a|0\rangle + b|1\rangle \in \mathcal{H}$ denote an arbitrary input state ($a, b \in \mathbf{C}, |a|^2 + |b|^2 = 1$). By linearity, the WZ machine acts on $|\phi\rangle$ as

$$|\phi\rangle_A|Q\rangle_Z \rightarrow |\Phi(\phi)\rangle_{ABX} = a|0\rangle_A|0\rangle_B|Q_0\rangle_X + b|1\rangle_A|1\rangle_B|Q_1\rangle_X.$$

Let us quantify the quality of the clones by the average fidelity $f_{WZ} = \int d\phi f_{WZ}(\phi)$, where $0 \leq f_{WZ}(\phi) \leq 1$ quantifies the overlap, or indistinguishability, between each output of the cloning machine and the state to clone;

$$f_{WZ}(\psi) = \langle\phi|Tr_{A,X}[|\Phi(\phi)\rangle\langle\Phi(\phi)|]|\phi\rangle = \langle\phi|Tr_{B,X}[|\Phi(\phi)\rangle\langle\Phi(\phi)|]|\phi\rangle.$$

One can check that the WZ machine achieves an average fidelity $f_{WZ} = 2/3$. We also note that the states are not all copied with the same fidelity.

Can we devise an imperfect cloning machine exhibiting a higher quality? Bužek and Hillery have brought an affirmative answer to this question [26]. They proposed the following unitary transformation:

$$|0\rangle_A|Q\rangle_Z \rightarrow \frac{1}{\sqrt{2}}(|00\rangle_{AB}|Q_{00}\rangle_X + \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)_{AB}|Q_{0+}\rangle_X), \tag{3.5}$$

$$|1\rangle_A|Q\rangle_Z \rightarrow \frac{1}{\sqrt{2}}(|11\rangle_{AB}|Q_{11}\rangle_X + \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)_{AB}|Q_{1+}\rangle_X). \tag{3.6}$$

This machine outputs two clones of equal fidelity $f_{BH} = 5/6$, and this fidelity is independent of the state to clone. It is instructive to write the cp-map associated with this transformation. In Schrödinger's picture, it reads:

$$\mathcal{C}_{BH} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H}^{\otimes 2}) : \rho \rightarrow \frac{2}{3} S_2(\rho \otimes \mathbf{1}) S_2, \tag{3.7}$$

where $S_2 = |00\rangle\langle 00| + |11\rangle\langle 11| + \frac{1}{2}(|01\rangle + |10\rangle)(\langle 01| + \langle 10|)$ is the projector onto the symmetric subspace of $\mathcal{H}^{\otimes 2}$: $\mathcal{H}_+^{\otimes 2}$. $\mathcal{C}_{BH}$then appears to have a very simple form, which allows for a physical interpretation. Cloning can be achieved by symmetrising the state of the qubit to be cloned with the completely random state of another qubit.

3.3 More quantum cloning machines

The no-cloning theorem and the BH cloning machine of qubits can be generalised in a variety of ways. One such generalisation is the case where one wants to produce M clones from N identical originals ($N < M$). One can also go beyond the case of qubits and consider the issue of cloning d-level systems. A lot of work has been devoted to such generalisations [27, 28, 18], providing a closed expression for the universal optimal symmetric $N \to M$ quantum cloning of d-level systems pure states [18]:

$$\mathcal{C} : \mathcal{B}(\mathcal{H}^{\otimes N}) \to \mathcal{B}(\mathcal{H}^{\otimes M}) : \rho \to \frac{d[N]}{d[M]} S_M(\rho \otimes \mathbf{1}^{\otimes M-N})S_M, \tag{3.8}$$

where $d[M] = \binom{d+M-1}{M}$ denotes the dimension of the symmetric subspace $\mathcal{H}_+^{\otimes M}$ and S_M denotes the projector onto $\mathcal{H}_+^{\otimes M}$. This machine achieves for each clone a fidelity

$$f_{opt}(d, N, M) = \frac{M - N + N(M + d)}{M(N + d)} \tag{3.9}$$

For $d = 2, N = 1, M = 2$, this machine reduces to the Bužek-Hillery machine, thus proving its optimality. The cp-map $\mathcal{C}$ is particularly interesting to discuss in three limit cases: $d \to \infty$, $N \to \infty$, and $M \to \infty$. The case $d \to \infty$ will be the object of the next chapter. Let us only discuss the two other cases here. When $N \to \infty$, Eq. (3.8) shows that the fidelity of the clones tends to unity. According to the no-cloning theorem, this means that the states to clone tend to become "more and more orthogonal", that is, the system to clone tend to become classical, which is obvious: $\forall |\psi\rangle, |\phi\rangle \in \mathcal{H}$,

$$\lim_{N\to\infty} |\langle\psi|\phi\rangle|^N = \begin{cases} 0 & \text{if } \psi \neq \phi \\ 1 & \text{if } \psi = \phi \end{cases}$$

In the limit $M \to \infty$, a strong link appears between quantum cloning and optimal state estimation [29]. Let $|\psi\rangle^{\otimes N} \in \mathcal{H}^{\otimes N}$ denote an unknown pure state. The limit $\lim_{M\to\infty} f(d, N, M)$ is the optimal fidelity that we can achieve when trying to estimate ψ. Indeed, it is easy to understand that

$$\lim_{M\to\infty} f(d, N, M) \leq f_{est}^{opt}(d, N), \tag{3.10}$$

where $f_{est}^{opt}(d, N)$ denotes the optimal fidelity achievable when trying to estimate from measurements performed on $|\psi\rangle^{\otimes N}$. This is so because when $M \to \infty$, measurements can be performed on each clone to perfectly determine their individual state $\rho_1(\psi)$, which is then read as an estimate for $|\psi\rangle\langle\psi|$. Interestingly, one can prove that the inequality (3.10) also goes the other way round [29]:

$$f_{est}^{opt}(d, N) \leq \lim_{M\to\infty} f(d, N, M). \tag{3.11}$$

Thus, in the limit $M \to \infty$, we can identify (the information yielded by) the output of an optimal $N \to M$ quantum cloning machine, and that of an optimal estimator (taking N replicas of the state to estimate as input).

Many variants of the machine (3.8) exist (see for example [30, 31, 32] and references therein). In the remainder of this chapter, we discuss one such generalisation.

3.4 Cloning of orthogonal qubits

As mentioned earlier, quantum cloning machines allow us to study how well quantum information can be distributed. As we shall see, the formalism of quantum cloning machines also allows to probe finely how well a quantum state can encode quantum information. Our starting point is a very interesting observation made by Gisin and Popescu [33]: the information about a direction in space is better encoded in two orthogonal qubits than in two identical ones. Consider the following situation. Suppose that an emitter, A, wants to communicate a direction (θ, ϕ) to a receiver, B, using just two qubits. We also suppose that A and B have previously agreed on a common reference frame. A can achieve this task by sending B two identical replicas of the state $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle$. Bob then performs some measurement on $|\psi\rangle^{\otimes 2}$ and gets an estimate (θ_e, ϕ_e). If the quality of this procedure is quantified by averaging (the square module of) the overlap between $|\psi\rangle$ and $|\psi_e\rangle = \cos\frac{\theta_e}{2}|0\rangle + e^{i\phi_e}\sin\frac{\theta_e}{2}|1\rangle$, then according to Eq. (3.9), Bob's fidelity for the guess is bounded by

$$\lim_{M\to\infty} f(2,2,M) \equiv F_{||} = 3/4. \tag{3.12}$$

The observation made by Gisin and Popescu is that A and B can apply a better strategy. Let $|\psi_\perp\rangle$ such that $\langle\psi|\psi_\perp\rangle = 0$. If A sends the state $|\psi\rangle|\psi_\perp\rangle$ instead of $|\psi\rangle^{\otimes 2}$, then there exists a povm on B's side which allows to achieve the fidelity

$$F_\perp = \frac{1}{2}(1 + \frac{1}{\sqrt{3}}) \approx 0.789.$$

Motivated by this result, we have considered the following question: Can M clones of a qubit $|\psi\rangle$ be produced from an orthogonal qubit pair $|\psi, \psi_\perp\rangle$ with a higher fidelity than from an identical pair $|\psi, \psi\rangle$? We here present a universal cloning machine acting on an orthogonal qubit pair that approximately implements the transformation $|\psi\rangle|\psi_\perp\rangle \to |\psi\rangle^{\otimes M}$ with the optimal fidelity [34]. For $M \geq 6$, this machine outperforms the corresponding $2 \to M$ cloning machine of qubits. We will also consider the possibility to realise this machine experimentally and propose a probabilistic implementation in quantum optics based on Parametric Down-Conversion. Our proposed setup is an extension of a scheme achieving "standard" cloning of qubits [35].

Let us first provide a simple argument on why we can expect the state $|\psi, \psi_\perp\rangle$ to be better cloned than $|\psi, \psi\rangle$. With an optimal measurement of $|\psi, \psi_\perp\rangle$, we can prepare M identical clones of $|\psi\rangle$, each with a fidelity $F_\perp$. In contrast, with a $2 \to M$ optimal universal cloning machine, we get, from Eq. (3.9), $F_{||}(M) = (3M+2)/(4M)$. Clearly, $F_{||}(M) < F_\perp$ for $M \geq 12$. Hence, this (non-optimal) measurement-based cloning of $|\psi, \psi_\perp\rangle$ is better than the standard cloning of $|\psi, \psi\rangle$ for sufficiently large values of M.

Optimal cloning of orthogonal qubits

Let us now seek for a unitary transformation which *optimally* approximates the transformation $|\psi\rangle|\psi_\perp\rangle \to |\psi\rangle^{\otimes M}$. Since we look for a transformation such that the final state of the clones is left invariant by permutations amongst them, we will suppose that the clones lie in the symmetric M-qubit space. Our motivation, when making this simplifying hypothesis, is that in the case of standard cloning, an optimal universal machine can always be chosen to be of this form [36]. Moreover, since the set of all states of the form $|\psi\rangle|\psi_\perp\rangle$ span the whole Hilbert space of two qubits, the most general transformation is of the form:

$$|i\rangle|j\rangle|R\rangle \to \sum_{k=0}^{M} |M,k\rangle|R_{ijk}\rangle, \qquad i,j = 0,1, \tag{3.13}$$

where $|R\rangle$ and $|R_{ijk}\rangle$ respectively denote the initial and final states of the ancilla, while $|M,k\rangle$ $(k=0,\ldots,M)$ denotes a symmetric M-qubit state with k qubits in state $|0\rangle$ and $M-k$ qubits in state $|1\rangle$.

The arbitrary state of a qubit $|\psi\rangle$ can be conveniently written as $|\psi\rangle = \pi(\Omega)|0\rangle = \sum_i \pi_{i0}(\omega)|i\rangle$, where the matrix $\pi(\Omega)$ is given by

$$\pi(\Omega) = \begin{pmatrix} \cos\frac{\theta}{2} & e^{-i\phi} - \sin\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{3.14}$$

with θ and ϕ denoting the usual polar and azimuthal angles pointing in direction Ω. The linearity of (3.13) implies that an arbitrary pair of orthogonal qubits transforms according to

$$|\psi\rangle|\psi_\perp\rangle \to |\Psi_{\rm out}(\psi)\rangle = \sum_{ijk} \pi_{i0}(\omega)\pi_{j1}(\Omega)|M,k\rangle|R_{ijk}\rangle. \tag{3.15}$$

We will measure the quality of the transformation by the average single-clone fidelity $F_\perp(M)$. Denoting by $\mathrm{Tr}_{1,\rm anc}$ the partial trace over the ancilla and all the clones but anyone, a straightforward calculation shows that

$$\begin{aligned} F_\perp(M) &= \int d\Omega \langle\psi|\,\mathrm{Tr}_{1,\rm anc}[|\Psi_{\rm out}(\psi)\rangle\langle\Psi_{\rm out}(\psi)|]\,|\psi\rangle \\ &= \sum_{i'j'k'}\sum_{ijk} \langle R_{i'j'k'}|R_{ijk}\rangle A_{ijk}^{i'j'k'}, \end{aligned} \tag{3.16}$$

where

$$A_{ijk}^{i'j'k'} = \sum_{n,n'} \langle n'|\mathrm{Tr}_1[\,|M,k\rangle\langle M,k'|\,]|n\rangle \int d\Omega\, \pi_{n0}(\Omega)\pi^*_{n'0}(\Omega)\pi_{i0}(\omega)\pi_{j1}(\Omega)\pi^*_{i'0}(\Omega)\pi^*_{j'1}(\Omega).$$

The coefficients $A_{ijk}^{i'j'k'}$ can be considered as matrix elements of an operator A acting on the space $\mathcal{H}\otimes\mathcal{H}_+^{\otimes M}$, where $\mathcal{H}$ denotes the Hilbert space of the two input qubits and $\mathcal{H}_+^{\otimes M}$ denotes the Hilbert space of symmetric states of M output qubits. Similarly, $\Gamma_{ijk}^{i'j'k'} = \langle R_{ijk}|R_{i'j'k'}\rangle$ define matrix elements of an operator Γ also acting on $\mathcal{H}\otimes\mathcal{H}_+^{\otimes M}$. The formula (3.16) for the fidelity thus simplifies to

$$F_\perp(M) = \mathrm{Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma A].$$

The operator Γ uniquely represents the completely positive cloning map, which transforms operators supported on $\mathcal{H}$ onto operators supported on $\mathcal{H}_+^{\otimes M}$. By definition, the operators A and Γ are Hermitian and positive semidefinite, $A \geq 0$ and $\Gamma \geq 0$.

Of course, the transformation (3.13) should be unitary, which reads

$$\sum_k \langle R_{i'j'k}|R_{ijk}\rangle = \delta_{i'i}\delta_{j'j}.$$

This is equivalent to

$$\mathrm{Tr}_{\mathcal{H}_+^{\otimes M}}[\Gamma] = \mathbf{1}_{\mathcal{H}}, \tag{3.17}$$

where $\mathbf{1}_{\mathcal{H}}$ is the identity operator on $\mathcal{H}$. Thus, introducing a set of Lagrange multipliers $\lambda_{ij}^{i'j'}$ for these unitarity constraints, our problem amounts to extremise the quantity

$W = \mathrm{Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[(A-\Lambda)\Gamma]$ under the constraint $\Gamma \geq 0$, where $\Lambda = \lambda \otimes S_M$ and λ is the matrix of Lagrange multipliers (S_M is the identity operator on $\mathcal{H}_+^{\otimes M}$). Varying W with respect to (the coefficients of) the eigenstates of the operator Γ (in a proper basis), we get the extremal equation

$$(A-\Lambda)\Gamma = 0 \tag{3.18}$$

for the optimal Γ. Following [37], this equation can be further transformed into a form suitable for numerical solution via repeated applications of

$$\Gamma = \Lambda^{-1} A\Gamma A \Lambda^{-1}, \qquad \lambda = (\mathrm{Tr}_{\mathcal{H}_+^{\otimes M}}[A\Gamma A])^{1/2}. \tag{3.19}$$

Note that the matrix $\lambda > 0$ is determined from the unitarity constraints.

By numerically solving Eq. (3.19) for $M = 2, \ldots 15$, we have guessed the general solution of Eq. (3.19). The transformation we obtain is:

$$|\psi, \psi_\perp\rangle \to \sum_{j=0}^{M} \alpha_{j,M} |(M-j)\psi, j\psi_\perp\rangle \otimes |(M-j)\psi_\perp, j\psi\rangle, \tag{3.20}$$

where

$$\alpha_{j,M} = (-1)^j \left[\frac{1}{\sqrt{2(M+1)}} + \frac{\sqrt{3}(M-2j)}{\sqrt{2M(M+1)(M+2)}} \right], \tag{3.21}$$

with $|j\psi, (M-j)\psi_\perp\rangle$ denoting a totally symmetric state of M qubits where j qubits are in the state $|\psi\rangle$ and $M-j$ qubits are in the state $|\psi_\perp\rangle$. The first M output qubits contain the clones of $|\psi\rangle$ while the other M qubits contain the clones of $|\psi_\perp\rangle$ (or anticlones). We shall prove below that the transformation Eq. (3.20) is indeed optimal.

First, we stress here that the cloning transformation (3.20) is unitary. Since this is not obvious from (3.20), let us present a proof of this. We can expand any state $|j\psi, (M-j)\psi_\perp\rangle$ in the basis $|M,k\rangle$ as

$$|j\psi, (M-j)\psi_\perp\rangle = \sum_{k=0}^{M} e^{i(j-k)\phi} D^M_{kj}(\theta) |M,k\rangle. \tag{3.22}$$

We will not need an explicit expression for the functions $D^M_{kj}(\theta)$ here, but we will only use some of their properties. Since the functions $D^M_{kj}(\theta)$ are elements of a (real) unitary matrix, they satisfy the orthogonality relation,

$$\sum_{j=0}^{M} D^M_{kj}(\theta) D^M_{lj}(\theta) = \delta_{kl}. \tag{3.23}$$

We will also use the following recurrence formula [38],

$$\begin{aligned} (2j-M) D^M_{kj}(\theta) \;\; = \;\; & (2k-M)\cos\theta \, D^M_{kj}(\theta) \\ & + \sin\theta \sqrt{(k+1)(M-k)} \, D^M_{k+1,j}(\theta) \\ & + \sin\theta \sqrt{k(M-k+1)} \, D^M_{k-1,j}(\theta). \end{aligned} \tag{3.24}$$

To prove that the transformation Eq.(3.20) is unitary, it is convenient to apply $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)^{\otimes M}$ on the last M qubits at the output of the cloner. Thus $|(M-j)\psi_\perp, j\psi\rangle \to$

$(-1)^j|(M-j)\bar{\psi}, j\bar{\psi}_\perp\rangle$ where $|\bar{\psi}\rangle = \sum_i \bar{\pi}_{i0}|i\rangle$. Next we expand $|(M-j)\psi, j\psi_\perp\rangle$ and $|(M-j)\bar{\psi}, j\bar{\psi}_\perp\rangle$ in the basis $|M,k\rangle$ using Eq. (3.22), and then utilise the recurrence formula (3.24). Finally, we can carry out the sum over j with the help of Eq. (3.23), resulting in

$$\begin{aligned}
|\Phi_{\rm out}(\psi)\rangle \quad = \quad & \sum_{k=0}^{M} [a_M + b_M(2k-M)] \cos^2\frac{\theta}{2} |M,k\rangle \otimes |M,k\rangle \\
& + \sum_{k=0}^{M} [a_M - b_M(2k-M)] \sin^2\frac{\theta}{2} |M,k\rangle \otimes |M,k\rangle \\
& + e^{i\phi} \sum_{k=0}^{M} b_M \sqrt{(M-k)(k+1)}\, \sin\theta |M,k\rangle \otimes |M,k+1\rangle \\
& + e^{-i\phi} \sum_{k=0}^{M} b_M \sqrt{k(M-k+1)}\, \sin\theta |M,k\rangle \otimes |M,k-1\rangle,
\end{aligned} \tag{3.25}$$

where the coefficients a_M and b_M read

$$a_M = \frac{1}{\sqrt{2(M+1)}}, \qquad b_M = \frac{\sqrt{3}}{\sqrt{2M(M+1)(M+2)}}.$$

The four terms on the right-hand side of Eq. (3.25) are proportional to the output states for the four input basis states $|01\rangle$, $|10\rangle$, $|00\rangle$, and $|11\rangle$, respectively. It is then easy to prove that the transformation $|\psi, \psi_\perp\rangle \to |\Phi_{\rm out}(\psi)\rangle$ preserves scalar products, hence is unitary.

Let us now consider the fidelity of the clones. We can see from Eq. (3.20) that the cloning machine preserves the symmetry of the input state $|\psi, \psi_\perp\rangle$, so the clones of both states $|\psi\rangle$ and $|\psi_\perp\rangle$ have the same fidelity. A little algebra shows this state-independent single-qubit fidelity can be obtained by summing a series:

$$F_\perp(M) = \sum_{j=0}^{M} \frac{M-j}{M}\, \alpha_{j,M}^2, \tag{3.26}$$

so that finally

$$F_\perp(M) = \frac{1}{2}\left(1 + \sqrt{\frac{M+2}{3M}}\right). \tag{3.27}$$

Upon comparing this fidelity to that of the optimal cloner for a pair of identical qubits $F_{||}(M)$, we see that $F_{||}(M) \geq F_\perp(M)$ for $M \leq 6$, while $F_\perp(M) > F_{||}(M)$ for $M > 6$ and the cloner (3.20) then outperforms the standard cloner. We also note that for $M \to \infty$, the fidelity $F_\perp(M)$ tends to the optimal measurement fidelity $F_\perp$, as expected.

To prove the optimality of our cloner, we invoke techniques adapted from the theory of semidefinite programming [39]. We observe that the trace of Lagrange multiplier λ provides an upper bound on the achievable fidelity $F_\perp(M) = {\rm Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma A]$. If $\lambda \otimes S_M - A \geq 0$ then it holds for any $\Gamma \geq 0$ that

$${\rm Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma\lambda \otimes S_M] \geq {\rm Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma A]. \tag{3.28}$$

It follows from the unitarity constraint Eq.(3.17) that $\mathrm{Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma\lambda\otimes S_M] = \mathrm{Tr}_{\mathcal{H}}[\lambda]$ does not depend on Γ. Thus $\mathrm{Tr}_{\mathcal{H}}[\lambda] \geq \mathrm{Tr}_{\mathcal{H},\mathcal{H}_+^{\otimes M}}[\Gamma A]$. From the numerical solution of Eqs. (3.19) we have in basis $|00\rangle$, $|11\rangle$, $|01\rangle$, $|10\rangle$,

$$\lambda = \frac{F_\perp(M)}{6}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}. \tag{3.29}$$

The block-diagonal matrix $\lambda\otimes S_M - A$ is positive semidefinite and has three different eigenvalues which read $\mu_1 = \frac{1}{12}\sqrt{\frac{M+2}{3M}}$, $\mu_2 = \frac{1}{3}\sqrt{\frac{M+2}{3M}}$, and $\mu_3 = 0$. Since the upper bound $\mathrm{Tr}_{\mathcal{H}}[\lambda] = F_\perp(M)$ is saturated by our cloning machine, we conclude that our cloner is optimal.

Implementation

We now propose a probabilistic implementation of the cloning transformation Eq.(3.20) via PDC. As we shall see, many technical difficulties would arise when achieving this implementation. Nevertheless, our main concern here is to stress that our cloner is, at least in principle, achievable. The experimental setup under consideration is shown in Fig.(3.4). This scheme is a straightforward extension of the setup suggested by Simon *et al.* [35] for conventional qubits, where the qubits are represented by the polarisation state of photons. We can identify $|0\rangle$ with vertical polarisation and $|1\rangle$ with horizontal polarisation states.

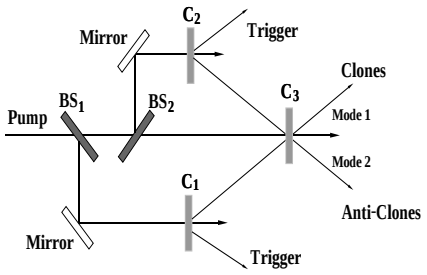

Figure 3.1: Setup for the cloning of orthogonal qubits via stimulated parametric down-conversion. For a detailed description, see text.

In optical PDC, a 'blue' photon can split into a pair of 'red' photons. (This occurs with a probability which is typically of order 10^{-3}.) These daughter photons are referred to as signal and idler, respectively. In our setup, three nonlinear crystals C_1, C_2, C_3 are pumped by a strong laser beam. In crystals C_1 and C_2, photons can be produced by pair, so we can verify the presence of signal photons by detecting the idler photons emerging from C_1 and C_2. If a single idler photon is detected in coincidence on each side, then we have one signal photon in each beam. The states of these two photons can be manipulated with the help of phase shifters and polarisation rotators in order to prepare the desired input state $|\psi,\psi_\perp\rangle$. The two photons then feed the signal and idler modes of a third nonlinear crystal C_3, where M clones are generated due to PDC.

In the limit of strong coherent pumping, the effective Hamiltonian describing the interaction in C_3 can be written as follows [35],

$$H = i\hbar g(a_{V1}^* a_{H2}^* - a_{H1}^* a_{V2}^*) + \text{h.c.}, \tag{3.30}$$

where a^*_{V1} and a^*_{H1} denote bosonic creation operators for photons in the first mode with vertical (V) or horizontal (H) polarisation, and similarly a^*_{V2} and a^*_{H2} are creation operators for photons in the second spatial mode. The constant g denotes the parametric gain. The time evolution is thus governed by the unitary transformation $U = \exp(-iHt/\hbar)$. With the help of the disentangling theorem [40], we can write the operator U in a factorised form

$$\begin{aligned} U &= e^{\Gamma a^*_{V1} a^*_{H2}} (\cosh\gamma)^{-(a^*_{V1}a_{V1} + a^*_{H2}a_{H2}+1)} e^{-\Gamma a_{V1}a_{H2}} \\ &\quad \times e^{-\Gamma a^*_{H1} a^*_{V2}} (\cosh\gamma)^{-(a^*_{H1}a_{H1} + a^*_{V2}a_{V2}+1)} e^{\Gamma a_{H1}a_{V2}}, \end{aligned}$$

where $\gamma = gt$ and $\Gamma = \tanh\gamma$. The Hamiltonian (3.30) has the important property of being invariant under general simultaneous SU(2) transformations on the polarisation vectors $(a_V,\, a_H)$ for modes 1 and 2 [35]. It is thus sufficient to consider the evolution of a basis state $|1\rangle_{V1}|0\rangle_{H1}|0\rangle_{V2}|1\rangle_{H2}$ (a single vertically polarised photon in mode 1 and a single horizontally polarised photon in mode 2) which represents the input state $|\psi, \psi_\perp\rangle \equiv |01\rangle$. Making use of the factorised form of U, we obtain the state at the output of the crystal C_3 in the form

$$\begin{aligned} \textstyle\sum_{M=0}^{\infty} \quad & \Gamma^{M-1}(1-\Gamma^2) \sum_{j=0}^{M} (-1)^j \left[(M-j)(1-\Gamma^2) - \Gamma^2\right] \\ \times \quad & |M-j\rangle_{V1}\, |j\rangle_{H1}\, |j\rangle_{V2}\, |M-j\rangle_{H2}, \end{aligned} \tag{3.31}$$

where $|k\rangle_l$ with $l = V1, H1, V2, H2$ denote the usual Fock states. For a fixed number M of photons in each mode 1 and 2, the output state (3.31) closely resembles the output state of our cloning machine (3.20) with the coefficients $\alpha_{j,M}(\Gamma) \approx \left[(M-j)(1-\Gamma^2) - \Gamma^2\right](-1)^j$. If we measure the number of photons in mode 2 and detect M photons, then we know that M photons representing M approximate clones of the input qubit $|\psi\rangle$ are present in mode 1. Note that the output of C_3 is not properly an M-qubit state but rather M indistinguishable photons distributed amongst two polarisation modes. Still, upon using an array of beam-splitters amongst M different modes, one can probabilistically obtain a proper M-qubit state.

In order to calculate the fidelity of these clones, we insert the properly normalised $\alpha_{j,M}(\Gamma)$ into the formula (3.26). We obtain

$$F(M,y) = \frac{3y^2 - 2y(2M+1) + \frac{3}{2}M(M+1)}{6y^2 - 6My + M(2M+1)} \tag{3.32}$$

where we have introduced $y \equiv \Gamma^2/(1-\Gamma^2) = \sinh^2\gamma$. The cloning fidelity thus depends on the parametric gain γ, so we must optimize this gain in order to achieve the highest possible fidelity. Upon solving $\frac{\partial F(M,y)}{\partial y} = 0$ for y, we find that, for a fixed value of M,

$$y_{\rm opt} = \frac{M}{2} - \frac{1}{2}\sqrt{\frac{M(M+2)}{3}}. \tag{3.33}$$

By inserting $y_{\rm opt}$ into Eq. (3.32), we recover the optimal fidelity (3.27). Furthermore, the postselected M-photon state at the output of the crystal C_3 coincides with the output of the cloning machine (3.20).

Even for a number of clones as low as 2, the amplification gain corresponding to y_{opt} is significantly larger than what is achievable with current technology. Moreover, there is a danger that the parametric approximation yielding the Hamiltonian (3.30) could fail for such large gains. Fortunately, one can easily verify that the function $F(M,y)$ is slowly varying with respect to y, that is, nearly optimal cloning devices can be achieved with realistic gains. For instance, $F(2,0) = 0.9$ instead of $F(2, y_{opt}) = 0.908$.

This approach of cloning based on PDC can be further extended to the approximate realization of the general cloning transformation

$$|\psi\rangle^{\otimes N}|\psi_\perp\rangle^{\otimes N'} \rightarrow |\psi\rangle^{\otimes M}.$$

For $N' = 1$, we have been able to derive the optimal fidelity for any N and $M \geq N$ by a similar calculation,

$$F_\perp(N, M) = \frac{N+1}{N+3} + \frac{3(N-1) + \sqrt{P/(N+2)}}{2M(N+3)} \tag{3.34}$$

with $P = (N-1)(N^2 - 15N - 18) + 8M(N+1)(M+3-N)$. It can be checked that there is again a value of M above which this cloner outperforms the standard $(N+1) \rightarrow M$ cloner. For large N, however, the advantage becomes marginal.

3.5 Summary

In summary, we have presented and discussed the no-cloning theorem, as well as transformations achieving approximate cloning. The study of such transformations tells us how information contained in one (or several) quantum system(s) can be distributed.

We have presented optimal quantum cloning machines taking orthogonal qubits as input, and shown how these machines can outperform standard quantum cloning machines. We have also shown how to implement them. We think we have thus contributed to bringing a better understanding of cloning. This study brings new evidence that, given a quantum system to encode some information, not all encoding schemes are equivalent. Actually, the best encoding to use depends on the information one wants to be able to extract. For example, a direction is better encoded in two orthogonal qubits, as we have seen. But if one wants to determine a direction orthogonal to the common direction of either two identical or two orthogonal qubits [1], then one would better use identical pairs [41]. An open question is to characterise (in general) the relation between the information one wants to optimally encode using a quantum system and the optimal way to prepare this quantum system.

[1] Two orthogonal qubits have the same direction on Bloch sphere.

Chapter 4

Quantum cloning of continuous variable systems

We address the issue of approximate cloning for continuous variable systems (or quantum oscillators). Gaussian cloning machines are presented, and the issue of phase-conjugation is investigated. Quantum cloning machines taking conjugate input are then studied and shown to outperform standard cloning machines in some cases.

4.1 Optimal cloning of Gaussian states

The Hilbert space associated with a quantum oscillator is $\mathcal{H} \equiv L^2(\mathbf{R})$, and is infinite-dimensional. Let us first consider what can we get from asking for universality in such a Hilbert space. Considering the limit for $d \to \infty$ of Eq. (3.9), we see that

$$\lim_{d\to\infty} f_{opt}(d, N, M) = \frac{N}{M}, \tag{4.1}$$

where N the number of input replicas, and $M > N$ the number of clones. In some circumstances such as quantum cryptography, one might want to study cloners which are optimal only for a subset $\mathcal{S} \subset \mathcal{H}$. Also, the fidelity is not always the most interesting figure of merit to consider. Can we then do better than Eq. (4.1)?

We will here concentrate on the situation where we only want to clone the set of minimum uncertainty coherent states: they satisfy

$$\langle \hat{x^2} - \hat{x}^2 \rangle = \langle \hat{p^2} - \hat{p}^2 \rangle = 1/2, \tag{4.2}$$

and can be parametrised as

$$\mathcal{S} = \{|\alpha\rangle : \alpha = \frac{1}{\sqrt{2}}(x + ip), x, p \in \mathbf{R}\}, \tag{4.3}$$

where $\langle\alpha|\hat{x}|\alpha\rangle = x$ and $\langle\alpha|\hat{p}|\alpha\rangle = p$. We will consider $N \to M$ symmetric Gaussian cloners (SGC). These cloners are linear, trace-preserving, completely positive maps $\mathcal{C}$ outputting M clones from $N \leq M$ identical replicas of an unknown coherent state $|\psi\rangle$. To simplify the analysis, we require that the joint state of the M clones $\mathcal{C}(|\psi^{\otimes N}\rangle\langle\psi^{\otimes N}|)$ is supported on the symmetric subspace of $\mathcal{H}^{\otimes M}$, and such that the partial trace over all output clones but (any) one is the bi-variate Gaussian mixture:

$$\begin{aligned} \rho_1(\psi) &= \mathrm{Tr}_{M-1}\mathcal{C}(|\psi^{\otimes N}\rangle\langle\psi^{\otimes N}|) \\ &= \frac{1}{\pi\sigma^2_{N,M}} \int d^2\beta\, e^{-|\beta|^2/\sigma^2_{N,M}}\, D(\beta)|\psi\rangle\langle\psi|D^*(\beta) \end{aligned} \tag{4.4}$$

where the integral is performed over all values of $\beta = (x+ip)/\sqrt{2}$ in the complex plane ($\hbar = 1$), and the operator

$$D(\beta) = \exp(\beta a^* - \bar{\beta} a)$$

achieves a displacement of x in position and p in momentum, with $\hat{a} = \frac{1}{\sqrt{2}}(\hat{x} + i\hat{p})$ and $a^* = \frac{1}{\sqrt{2}}(\hat{x} - i\hat{p})$ denoting the annihilation and creation operators, respectively[1]. Thus, the copies yielded by a SGC are affected by an equal Gaussian noise $\sigma_x^2 = \sigma_p^2 = \sigma_{N,M}^2$ on the conjugate variables x and p. (It will turn out that the resulting cloning fidelity $f_g = \langle\psi|\rho_1(\psi)|\psi\rangle$ is invariant for all coherent states of $\hat{x}$ and $\hat{p}$.) We will show that a lower bound on the noise variance $\sigma_{N,M}^2$ is given by

$$\overline{\sigma}_{N,M}^2 = \frac{1}{N} - \frac{1}{M} \tag{4.5}$$

implying in turn that the optimal cloning fidelity for Gaussian cloning of coherent states is bounded by

$$f_{N,M} = \frac{MN}{MN + M - N} \tag{4.6}$$

Let us first prove (4.5) in the case $(N, M) = (1, 2)$. This case is interesting to single out because it demonstrates the link between quantum cloning and the problem of simultaneously measuring a pair of conjugate observables on a single quantum system. Our starting point is thus the Arthurs and Kelly relation [42] constraining any attempt to measure $\hat{x}$ and $\hat{p}$ simultaneously on a quantum system:

$$\sigma_x^2(1)\, \sigma_p^2(1) \geq 1, \tag{4.7}$$

where $\sigma_x^2(1)$ and $\sigma_p^2(1)$ denote the variance of the *measured* values of $\hat{x}$ and $\hat{p}$, respectively, when simultaneously measuring $\hat{x}$ and $\hat{p}$ on some quantum state ρ.

It is crucial to clearly distinguish between the Arthurs and Kelly relation (4.7), and the Heisenberg uncertainty relation:

$$\delta\hat{x}^2 \delta\hat{p}^2 \geq 1/4, \tag{4.8}$$

where $\delta\hat{x}^2$ (resp. $\delta\hat{p}^2$) are *intrinsic* variance of the observable $\hat{x}$ (resp. $\hat{p}$)for any quantum state ρ. The Heisenberg relation is valid independently from any measurement performed on the state ρ. It precisely answers the question:

> "For a quantum system prepared in the state ρ, to what extent can we simultaneously *define* (or assign values) to both the observables $\hat{x}$ and $\hat{p}$?"

In particular, the Heisenberg relation holds even if we have a perfect knowledge of the state ρ. In contrast, the Arthurs-Kelly relation quantifies the trade-off between the information about $\hat{x}$ and the information about $\hat{p}$, that one can acquire during a single *measurement* on the state ρ.

So, the best possible simultaneous measurement of $\hat{x}$ and $\hat{p}$ with a same precision satisfies $\sigma_x^2(1) = \sigma_p^2(1) = 1$. Compared with the intrinsic noise of a minimum-uncertainty wave packet $\delta\hat{x}^2 = \delta\hat{p}^2 = 1/2$, we see that the joint measurement of x and p effects an additional noise of minimum variance $1/2$. Now, let a coherent state $|\alpha\rangle$ be processed by a $1 \to 2$ SGC, and let $\hat{x}$ be measured at one output of the cloner while

[1] N.B. In the remainder of this thesis, we will sometimes omit the hats on operators when the context is clear.

$\hat{p}$ is measured at the other output. As cloning should obey inequality (4.7), we must have

$$\Delta\hat{x}^2\ \Delta\hat{p}^2 \geq 1 \tag{4.9}$$

where $\Delta\hat{x}^2$ (resp. $\Delta\hat{p}^2$) refers to the intrinsic variance of the observable $\hat{x}$ (resp. $\hat{p}$) for the state $\rho_1(\alpha)$. Using Eq. (4.4), we get

$$(\delta\hat{x}^2 + \sigma_{1,2}^2)(\delta\hat{p}^2 + \sigma_{1,2}^2) \geq 1 \tag{4.10}$$

Now upon using (4.8), and the relation

$$a^2 + b^2 \geq 2\sqrt{a^2b^2}\ \forall a, b \in \mathbf{R},$$

we conclude that the noise variance is constrained by

$$\sigma_{1,2}^2 \geq \overline{\sigma}_{1,2}^2 = 1/2, \tag{4.11}$$

thus verifying Eq. (4.5) in the case $(N, M) = (1, 2)$.

A similar argument can be used to characterise the output copies of an asymmetric quantum cloning machine, where the qualities of the clones are not identical and where one might desire that the added noise due to cloning is different for both quadratures. Using, Eq.(4.7), one easily shows that the following relations hold:

$$\sigma_{x,1}^2\sigma_{p,2}^2 \geq 1/4 \tag{4.12}$$

$$\sigma_{p,1}^2\sigma_{x,2}^2 \geq 1/4 \tag{4.13}$$

where $\sigma_{x,1}^2$ (resp. $\sigma_{p,1}^2$) refers to the added x quadrature (resp. p quadrature) added noise for the first clone, and where $\sigma_{x,2}^2$ and $\sigma_{p,2}^2$ are defined likewise. The relations (4.12) will be useful when discussing quantum cryptography.

Let us now prove Eq. (4.5) in the general case. Our proof is connected to quantum state estimation theory similarly to what was done for quantum bits in [**?**]. The key idea is that cloning should not be a way of circumventing the noise limitation encountered in any measuring process. More specifically, our bound relies on the fact that cascading an $N \to M$ cloner with an $M \to L$ cloner results in a $N \to L$ cloner which cannot be better that the *optimal* $N \to L$ cloner. We make use of the property that cascading two SGCs results in a single SGC whose variance is simply the sum of the variances of the two component SGCs (see below). Hence, the variance $\overline{\sigma}_{N,L}^2$ of the *optimal* $N \to L$ SGC must satisfy

$$\overline{\sigma}_{N,L}^2 \leq \sigma_{N,M}^2 + \sigma_{M,L}^2. \tag{4.14}$$

In particular, if the $M \to L$ cloner is itself optimal and $L \to \infty$,

$$\overline{\sigma}_{N,\infty}^2 \leq \sigma_{N,M}^2 + \overline{\sigma}_{M,\infty}^2 \tag{4.15}$$

As discussed in Chapter 3, in the limit $M \to \infty$, estimators and quantum cloning machines tend to become essentially identical devices. Thus Eq. (4.15) means that cloning the N replicas of a system before measuring the M resulting clones does not provide a mean to enhance the accuracy of a direct measurement of the N replicas.

Let us now estimate $\overline{\sigma}_{N,\infty}^2$, that is, the variance of an optimal joint measurement of $\hat{x}$ and $\hat{p}$ on N replicas of a system. From quantum estimation theory [43], we know that the variance of the measured values of $\hat{x}$ and $\hat{p}$ on a single system, respectively $\sigma_x^2(1)$ and $\sigma_p^2(1)$, are constrained by

$$g_x\sigma_x^2(1) + g_p\sigma_p^2(1) \geq g_x\delta\hat{x}^2 + g_p\delta\hat{p}^2 + \sqrt{g_xg_p} \tag{4.16}$$

for all values of the constants $g_x, g_p > 0$. Note that, for each value of g_x and g_p, a specific povm based on a resolution of identity in terms of squeezed states, whose squeezing Δ is a function of g_x. and g_p, achieves this bound (see [43]). Squeezed states satisfy $\langle \hat{x^2} - \hat{x}^2 \rangle = \Delta^2$ and $\langle \hat{p^2} - \hat{p}^2 \rangle = 1/4\Delta^2$. Moreover, when a measurement is performed on N independent and identical systems, the r. h. s. of (4.16) is reduced by a factor N^{-1}, as in classical statistics [44]. So, applying N times the optimal single-system povm is the best joint measurement when N replicas are available since it yields $\sigma_x^2(N) = N^{-1}\sigma_x^2(1)$ and $\sigma_p^2(N) = N^{-1}\sigma_p^2(1)$. Hence, using Eq. (4.16) for a coherent state ($\delta\hat{x}^2 = \delta\hat{p}^2 = 1/2$) and requiring $\sigma_x^2(N) = \sigma_p^2(N)$, the tightest bound is obtained for $g_x = g_p$. It yields $\overline{\sigma}_{N,\infty}^2 = 1/N$, which, combined with Eq. (4.15), gives the minimum noise variance induced by cloning, Eq. (4.5).

It now only remains to prove the validity of Eq. (4.15), that is, that the variance of two cascaded SGCs add. Consider an $N \to M$ SGC, followed by a $M \to L$ SGC. Let ρ be an arbitrary density operator supported on $\mathcal{H}^{\otimes M}$. Since it is self-adjoint and compact, ρ has a denumerable spectrum: it can be expanded as $\rho = \sum_{i=1}^{\infty} \lambda_i |\xi_i\rangle\langle\xi_i|$ with $\langle\xi_i|\xi_j\rangle = \delta_{ij}$, $\lambda_i \geq 0$ and $\sum_{i=1}^{\infty} \lambda_i = 1$. Note that $\forall \epsilon > 0$, $\exists d$ such that $|\sum_{i=1}^{d} \lambda_i - 1| < \epsilon$. Therefore, the output of the first cloner can be decomposed as $\rho_M = \rho_d + \epsilon_d\, B_d$ where $\rho_d = \sum_{i=1}^{d} \lambda_i |\xi_i\rangle\langle\xi_i|$ is supported on a d-dimensional subspace of $\mathcal{H}^{\otimes M}$, B_d is a bounded operator, and $\lim_{d\to\infty} \epsilon_d = 0$. Since ρ_M belongs to the symmetric subspace of $\mathcal{H}^{\otimes M}$, so will ρ_d. Hence, we know that we can write ρ_d in the form of a pseudo-mixture of pure product states $\rho_d = \sum_{i=1}^{d} \alpha_i |\phi_i^{\otimes M}\rangle\langle\phi_i^{\otimes M}|$ where the coefficients α_i are not necessarily positive but satisfy $\sum_{i=1}^{d} \alpha_i = 1$ (see [18, 29]). Thus, when cloning a state $|\psi^{\otimes N}\rangle$, we have

$$C_{N,M}(|\psi^{\otimes N}\rangle\langle\psi^{\otimes N}|) = \sum_{i=1}^{d} \alpha_i |\phi_i^{\otimes M}\rangle\langle\phi_i^{\otimes M}| + \epsilon_d\, B_d \tag{4.17}$$

Then, since the cloning map $C_{N,M}$ is linear, cascading the two cloners yields

$$C_{M,L}C_{N,M}(|\psi^{\otimes N}\rangle\langle\psi^{\otimes N}|) = \sum_i \alpha_i C_{M,L}(|\phi_i^{\otimes M}\rangle\langle\phi_i^{\otimes M}|) + \epsilon_d C_{M,L}(B_d).$$

As this expression is a density operator (thus bounded) and the first term of its r.h.s. is positive, $C_{ML}(B_d)$ must be bounded. Thus, the second term of the r. h. s. of Eq. (4.17) becomes negligible when $d \to \infty$. Now, using Eq.(4.4), we have

$$\begin{aligned} \mathrm{Tr}_{L-1} C_{M,L}C_{N,M}(|\psi^{\otimes N}\rangle\langle\psi^{\otimes N}|) = \int d^2\gamma\, d^2\beta\, \mathrm{e}^{-\frac{|\gamma|^2}{2\sigma_{M,L}^2}} \\ \mathrm{e}^{-\frac{|\beta|^2}{2\sigma_{N,M}^2}}\, D(\gamma+\beta)|\psi\rangle\langle\psi|D^*(\gamma+\beta) + O(\eta_d) \end{aligned} \tag{4.18}$$

with $\lim_{d\to\infty} \eta_d = 0$. A little algebra then shows that this last expression is a Gaussian mixture, centred about the original state, whose variance is $\sigma_{M,L}^2 + \sigma_{N,M}^2$.

It is now easy to compute the fidelity of the optimal $N \to M$ SGC when a coherent state $|\alpha\rangle$ is copied. Using Eq. (4.4) and the identity $|\langle\alpha|\alpha'\rangle|^2 = \exp(-|\alpha - \alpha'|^2)$, we obtain

$$f_{N,M} = \langle\alpha|\rho_1|\alpha\rangle = \frac{1}{1 + \overline{\sigma}_{N,M}^2} \tag{4.19}$$

which results in Eq. (4.6). As expected, all coherent states are copied with a same fidelity. (Note, however, that this property does not extend to all states of $\mathcal{H}$.) Equations (4.5) and (4.6) are consistent with the work of Cerf-Ipe-Rottenberg [45], where

a $1 \to 2$ cloner of Gaussian states was derived but not proven to be optimal. They also fulfill the natural requirement that the cloning fidelity increases with the number of input replicas. For instance, considering a $kN \to kM$ SGC with a positive integer k, we find that $\frac{\Delta\overline{\sigma}^2_{N,M}}{\Delta k} < 0$ (and $\frac{\Delta f}{\Delta k} > 0$). At the limit $N \to \infty$, we have $f \to 1$, $\forall M$, that is, classical copying is allowed. Finally, for $M \to \infty$, that is, for an optimal measurement, we get $f \to N/(N+1)$. In particular, it implies that the best simultaneous measurement of $\hat{x}$ and $\hat{p}$ on a single system gives a fidelity $1/2$, a well-known result.

It is worth noting that optimally cloning squeezed states requires a variant of these SGCs. Let us consider for instance a family of quadrature squeezed states with squeezing parameter r. For such a family, the best symmetric cloner must have the form of Eq. (4.4), but using the definition $\beta = (\frac{x}{\sigma} + i\sigma p)/\sqrt{2}$ with $\sigma = \exp(r)$. These cloners naturally generalize the SGCs and give the same cloning fidelity, Eq. (4.6), for those squeezed states.

4.2 Implementation of Gaussian quantum cloning machines

Now that we have derived upper bounds on optimal cloning, we will show that these bounds are achievable, and exhibit an explicit optimal cloning transformation. Remarkably, implementing this transformation only requires a phase-insensitive linear amplifier and a network of beam splitters. An experimental demonstration of this continuous cloner should therefore be in the scope of current technology. We will also discuss the link between the issue of optimal quantum cloning and that of the optimal amplification of quantum states.

We will first state what we expect from a quantum cloning machine. Again, we will start by considering the special case of duplication before treating the general case of N original replicas and $M \geq N$ output clones.

Let $|\Psi\rangle = |\alpha\rangle^{\otimes N} \otimes |0\rangle^{\otimes M-N} \otimes |0\rangle_z$ denote the initial joint state of the N input modes to be cloned (prepared in the coherent state $|\alpha\rangle$), the additional $M-N$ blank modes, and an ancillary mode z. The blank modes and the ancilla are assumed to be initally in the vacuum state $|0\rangle$. Let $\{x_k, p_k\}$ denote the pair of quadrature operators associated with each mode k involved by the cloning transformation, $k = 0 \dots N-1$ refers to the N original input modes, and $k = N \dots M-1$ refers to the additional blank modes. Cloning can be thought of as some unitary transformation

$$U : \mathcal{H}^{\otimes M} \to \mathcal{H}^{\otimes M} : |\Psi\rangle \to U|\Psi\rangle = |\Psi''\rangle$$

Alternatively, in Heisenberg picture, this transformation can be described by a canonical transformation of the operators $\{x_k, p_k\}$:

$$x''_k = U^* \, x_k \, U, p''_k = U^* \, p_k \, U, \tag{4.20}$$

while leaving the state $|\Psi\rangle$ invariant. We will work here in Heisenberg picture because cloning turns out to be much simpler to study from that point of view. We will now impose several requirements on the transformation Eq. (4.20) that translate the expected properties for an optimal cloning transformation.

First, we require that the M output modes quadratures to have the same mean values as the the input mode:

$$\langle {x_k}'' \rangle = \langle \psi | x_0 | \psi \rangle, \quad k = 0 \dots M-1, \tag{4.21}$$

$$\langle {p_k}'' \rangle = \langle \psi | p_0 | \psi \rangle, \quad k = 0 \dots M-1. \tag{4.22}$$

This means that the state of the clones is centred on the original coherent state. Our second requirement is covariance with respect to rotation in phase space. Coherent states have the property that quadrature variances are left invariant by complex rotations in phase space. So, for any mode k involved in the cloning process and for any operator $v_k = cx_k + dp_k$ (where c, d are complex numbers satisfying $|c|^2 + |d|^2 = 1$), the error variance $\sigma^2_{v_k}$ is the same:

$$\sigma^2_{v_k} = \langle (v_k)^2 \rangle - \langle v_k \rangle^2 = \Delta x^2_{vac} = 1/2.$$

We impose this property to be conserved through the cloning process. Taking optimality into account, Eq. (4.5), rotation covariance yields:

$$\sigma^2_{v''_k} = (1 + \frac{2}{N} - \frac{2}{M})\Delta x^2_{vac}, \tag{4.23}$$

where $v''_k = cx''_k + dp''_k$.

Our third requirement is, of course, the unitarity of the transformation. In Heisenberg picture, unitarity is equivalent to demanding that the commutation rules are conserved through the evolution [46]:

$$[x_j{}', x_k{}'] = [p_j{}', p_k{}'] = 0, \ [x_j{}', p_k{}'] = i\delta_{jk}. \tag{4.24}$$

Let us first focus on duplication ($N = 1, M = 2$). A simple transformation meeting the three conditions mentioned above is given by:

$$\begin{aligned}
x''_0 &= x_0 + \frac{x_1}{\sqrt{2}} + \frac{x_z}{\sqrt{2}}, \qquad p''_0 = p_0 + \frac{p_1}{\sqrt{2}} - \frac{p_z}{\sqrt{2}}, \\
x''_1 &= x_0 - \frac{x_1}{\sqrt{2}} + \frac{x_z}{\sqrt{2}}, \qquad p''_1 = p_0 - \frac{p_1}{\sqrt{2}} - \frac{p_z}{\sqrt{2}}, \\
x'_z &= x_0 + \sqrt{2}\, x_z, \qquad\qquad p'_z = -p_o + \sqrt{2}\, p_z.
\end{aligned} \tag{4.25}$$

This transformation clearly conserves the commutation rules, and yields the expected mean values ($\langle x_0 \rangle, \langle p_0 \rangle$) for the two clones (modes $0''$ and $1''$). Also, one can check that the quadrature variances of both clones are equal to $2\Delta x^2_{vac}$, in accordance with Eq.(4.23). This transformation actually coincides with the Gaussian cloning machine introduced by Cerf et al. [45]. Interestingly, we note here that the state in which the ancilla z is left after cloning is centered on $(x_0, -p_0)$, that is the *phase-conjugated* state $|\bar{\alpha}\rangle$. This means that, in analogy with the universal qubit cloning machine [26], the continuous-variable cloner generates an "anticlone" (or time-reversed state) together with the two clones.

Now, let us show how this duplicator can be implemented in practice. Equation (4.25) can be interpreted as a sequence of two canonical transformations:

$$\begin{aligned}
a'_0 &= \sqrt{2}a_0 + a^*_z, \qquad a'_z = a^*_0 + \sqrt{2}a_z, \\
a''_0 &= \frac{1}{\sqrt{2}}(a'_0 + a_1), \quad a''_1 = \frac{1}{\sqrt{2}}(a'_0 - a_1).
\end{aligned} \tag{4.26}$$

As shown in Fig. 4.2, the interpretation of this transformation becomes then straightforward: the first step (which transforms a_0 and a_z into a'_0 and a'_z) is a phase-insensitive amplifier whose (power) gain G is equal to 2, while the second step (which transforms a'_0 and a_1 into a''_0 and a''_1) is a phase-free 50:50 beam splitter. Clearly, rotational covariance is guaranteed here by the use of a *phase-insensitive* amplifier. As discussed in [46], the ancilla z involved in linear amplification can always be chosen such that $\langle a_z \rangle = 0$, so that we have $\langle a''_0 \rangle = \langle a''_1 \rangle = \langle a_0 \rangle$ as required. Finally, the optimality of

our cloner can be confirmed from known results on linear amplifiers. For an amplifier of gain G, the quadrature variances of a_z are bounded by [46]:

$$\sigma^2_{a_z} \geq (G-1)/2. \tag{4.27}$$

Hence, the optimal amplifier of gain $G = 2$ yields $\sigma^2_{a_z} = 1/2$, so that our cloning transformation is optimal.

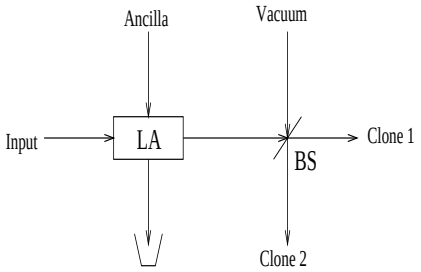

Figure 4.1: Implementation of a $1 \rightarrow 2$ cloning machine. LA stands for linear amplifier, and BS represents a balanced beam splitter.

Let us now derive an $N \rightarrow M$ cloning transformation. To achieve cloning, energy has to be brought to each of the $M - N$ blank modes in order to drive them from the vacuum state to a state which has the desired mean value. We will again achieve this operation with the help of a linear amplifier. From Eq.(4.27), we see that the cloning induced noise essentially originates from the amplification process, and grows with the gain of amplifier. So, we will preferably amplify as less as possible. Loosely speaking, the cloning procedure should then be as follows: (i) symmetrically amplifying the N input modes by *concentrating* them into one single mode, which is then amplified; (ii) symmetrically *distributing* the output of this amplifier amongst the M output modes. As we will see, a convenient way to achieve these concentration and distribution processes is provided by the Discrete Fourier Transform (DFT). Cloning is then achieved by the following three-step procedure (see Fig. 4.2). First step: a DFT (acting on N modes),

$$a'_k = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} \exp(ikl2\pi/N)\, a_l, \tag{4.28}$$

with $k = 0 \ldots N-1$. This operation concentrates the energy of the N input modes into one single mode (renamed a_0) and leaves the remaining $N-1$ modes ($a'_1 \ldots a'_{N-1}$) in the vacuum state. Second step: the mode a_0 is amplified with a linear amplifier of gain $G = M/N$. This results in

$$\begin{aligned} a'_0 &= \sqrt{\frac{M}{N}}\, a_0 + \sqrt{\frac{M}{N} - 1}\, a^*_z, \\ a'_z &= \sqrt{\frac{M}{N} - 1}\, a^*_0 + \sqrt{\frac{M}{N}}\, a_z. \end{aligned} \tag{4.29}$$

Third step: amplitude distribution by performing a DFT (acting on M modes) between the mode a'_0 and $M-1$ modes in the vacuum state:

$$a''_k = \frac{1}{\sqrt{M}} \sum_{l=0}^{M-1} \exp(ikl2\pi/M)\, a'_l, \tag{4.30}$$

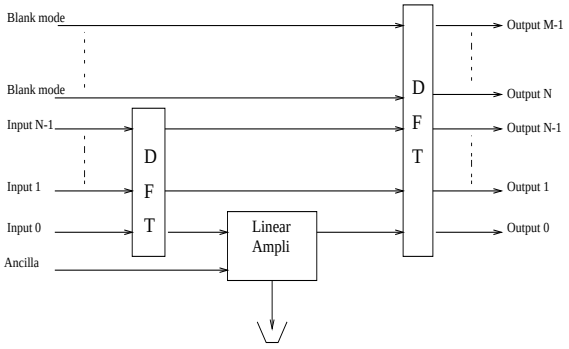

Figure 4.2: Implementation of an $N \to M$ cloning machine.

with $k = 0 \ldots M-1$, and $a'_i = a_i$ for $i = N \ldots M-1$. The DFT now distributes the energy contained in the output of the amplifier amongst the M output clones.

It is readily checked that this procedure meets our three requirements, and is optimal provided that the amplifier is optimal, that is $\sigma^2_{a_z} = [(M/N) - 1]/2$. The quadrature variances of the M output modes coincide with Eq. (4.5). As in the case of duplication, the quality of cloning decreases as $\sigma^2_{a_z}$ increases, that is cloning and amplifying coherent states are two equivalent problems. For $1 \to 2$ cloning, we have seen that the final amplitude distribution amongst the output clones is achieved with a single beam splitter. In fact, any unitary matrix such as the DFT used here can be realised with a sequence of beam splitters (and phase shifters) (see [47] and references therein). This means that the $N \to M$ cloning transformation can be implemented using only passive elements except for a single linear amplifier.

We will now explicitly give the *simplest* beam splitter combination that enables the above transformation. For convenience, let us now use the indices $k = 1 \ldots N$ for the N original input modes a_k, and $k = N+1 \ldots M$ for the additional blank modes a_k. With an ideal (phase-free) beam splitter operation acting on two modes c_k and c_l,

$$\begin{pmatrix} c'_k \\ c'_l \end{pmatrix} = \begin{pmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \begin{pmatrix} c_k \\ c_l \end{pmatrix} , \tag{4.31}$$

we define a matrix $B_{kl}(\theta)$ which is an M-dimensional identity matrix with the entries I_{kk}, I_{kl}, I_{lk}, and I_{ll} replaced by the corresponding entries of the above beam splitter matrix. Now we can define a sequence of beam splitters acting on M modes ("M-splitter") as

$$\begin{aligned} \mathcal{U}(M) \quad \equiv \quad & B_{M-1\,M}\left(\sin^{-1}\frac{1}{\sqrt{2}}\right) B_{M-2\,M-1}\left(\sin^{-1}\frac{1}{\sqrt{3}}\right) \\ & \times \cdots \times B_{12}\left(\sin^{-1}\frac{1}{\sqrt{M}}\right) . \end{aligned} \tag{4.32}$$

The individual beam splitters in Eq. (4.32) depend only on their reflectivity/transmittance parameter θ. In order to concentrate the N identical inputs, we send them now through an inverse N-splitter,

$$\begin{pmatrix} a'_1 & a'_2 & \cdots & a'_N \end{pmatrix}^T = \mathcal{U}^*(N) \begin{pmatrix} a_1 & a_2 & \cdots & a_N \end{pmatrix}^T . \tag{4.33}$$

Again, we end up with one mode (renamed a_1) having non-zero mean value and $N-1$ modes ($a'_2 \ldots a'_N$) in the vacuum state. After amplifying mode a_1, $a'_1 = \sqrt{M/N}\, a_1 +$

$\sqrt{M/N-1}\,a_z^*$, etc., a final M-splitter operation yields the output clones:

$$\left(\begin{array}{cccc} a_1'' & a_2'' & \cdots & a_M'' \end{array}\right)^T = \mathcal{U}(M) \left(\begin{array}{cccc} a_1' & a_2' & \cdots & a_M' \end{array}\right)^T , \tag{4.34}$$

with $a_i' = a_i$ for $i = N+1 \ldots M$.

Since the amplification produces extra noise, our cloning circuits used as little amplification as possible. However, rather surprisingly, by first amplifying each input copy $k = 1 \ldots N$ individually,

$$\begin{aligned} a_k' &= \sqrt{\frac{M}{N}}\, a_k + \sqrt{\frac{M}{N}-1}\, a_{z,k}^*, \\ a_{z,k}' &= \sqrt{\frac{M}{N}-1}\, a_k^* + \sqrt{\frac{M}{N}}\, a_{z,k}, \end{aligned} \tag{4.35}$$

a circuit can also be constructed that yields optimum fidelities. In the next step, the amplified modes are *each* sent together with $M-1$ vacuum modes $b_{k,1}$, $b_{k,2}$,...,$b_{k,M-1}$ through an M-splitter

$$\begin{aligned} &\left(\begin{array}{cccc} a_{k,1}' & a_{k,2}' & \cdots & a_{k,M}' \end{array}\right)^T = \\ &\quad \mathcal{U}(M) \left(\begin{array}{cccc} a_k' & b_{k,1} & \cdots & b_{k,M-1} \end{array}\right)^T . \end{aligned} \tag{4.36}$$

The NM output modes after this operation can be written as

$$a_{k,l}' = \frac{1}{\sqrt{N}} a_k + \sqrt{\frac{M-N}{MN}} a_{z,k}^* + d_{k,l} \ , \tag{4.37}$$

where $l = 1 \ldots M$. The noise in each M-splitter output coming from the $M-1$ vacuum inputs is represented by mode $d_{k,l}$ having zero mean value and quadrature variances of $(M-1)/2M$. The final step now consists of M inverse N-splitters acting on all modes with the same index l, i.e., the N modes $a_{k,1}'$, and the N modes $a_{k,2}'$, etc. The output modes at each N-splitter,

$$\begin{aligned} &\left(\begin{array}{cccc} a_l'' & e_{1,l} & \cdots & e_{N-1,l} \end{array}\right)^T = \\ &\quad \mathcal{U}^*(N) \left(\begin{array}{cccc} a_{1,l}' & a_{2,l}' & \cdots & a_{N,l}' \end{array}\right)^T , \end{aligned} \tag{4.38}$$

contain only noise except for one mode,

$$a_l'' = \sum_{k=1}^{N} \left(\frac{1}{N} a_k + \sqrt{\frac{M-N}{MN^2}} a_{z,k}^* + \frac{1}{\sqrt{N}} d_{k,l} \right) \ . \tag{4.39}$$

Again, all M clones are optimal, although additional noise has been introduced at the intermediate steps which results in $M(N-1)$ "waste" output modes. However, this particular circuit points out that $N \to M$ cloning of coherent states is effectively a "classical plumbing" procedure distributing classical amplitudes.

Finally, we note that for squeezed-state inputs rather than coherent states, the transformations and circuits presented require all auxiliary vacuum modes (the blank modes and the ancillary mode z) be correspondingly squeezed in order to maintain optimum cloning fidelities. This means, in particular, that the amplifier mode z needs to be controlled which requires a device different from a simple phase-insensitive amplifier, namely a two-mode parametric amplifier. One can say that the cloning machine capable of optimum cloning of all squeezed states with *fixed* and *known* squeezing then operates in a non-universal fashion with respect to all possible squeezed states at the input [13].

4.3 Phase-Conjugation

Before studying phase-conjugate input quantum cloning machines, we want to have a look at phase-conjugation and at the information content of a pair of phase-conjugate quantum coherent states. This work was stimulated by a paper of Gisin and Popescu [33], where similar questions were considered for qubits. Interestingly, our results are qualitatively similar to theirs, showing once more the relative analogy between quantum information processing with quantum bits and with continuous variables [48].

The phase conjugation operation consists in flipping the sign of the quadrature $\hat{p}$ while keeping the quadrature $\hat{x}$ unchanged, that is, replacing $\hat{a}$ by its Hermitian conjugate $\hat{a}^*$. Clearly, this operation is impossible as it does not conserve the commutation relation: if $\hat{b} = \hat{a}^*$ is the resulting mode, we have $[\hat{b}, \hat{b}^*] = -[\hat{a}, \hat{a}^*] = -1$ instead of 1 ($\hbar = 1$).

A heuristic argument can be used to show that this operation cannot be performed with an added noise that is lower than a minimum equal to twice the vacuum noise. Let us consider two modes (mode 0 and 1) that are initially prepared in the Einstein-Podolsky-Rosen (EPR) state [49], that is, the common eigenstate of $\hat{X} = \hat{x}_0 - \hat{x}_1$ and $\hat{P} = \hat{p}_0 + \hat{p}_1$ with zero eigenvalue for both operators $\hat{X}$ and $\hat{P}$. Since $[\hat{X}, \hat{P}] = 0$, these operators can be diagonalised simultaneously, so that the EPR state can be understood as representing two particles with a relative position $x_0 - x_1$ and a total momentum $p_0 + p_1$ both arbitrarily close to zero. Assume now that we apply a perfect phase conjugation on mode 1, that is, $\hat{x}_1' = \hat{x}_1$ and $\hat{p}_1' = -\hat{p}_1$, while mode 0 is left unchanged. The EPR state is then transformed into the common eigenstate with zero eigenvalue of the operators

$$\begin{aligned} \hat{X}' &= \hat{x}_0' - \hat{x}_1', \\ \hat{P}' &= \hat{p}_0' - \hat{p}_1'. \end{aligned} \tag{4.40}$$

We thus expect that $[\hat{X}', \hat{P}'] = 0$ since $\hat{X}' = \hat{X}$ and $\hat{P}' = \hat{P}$. However, $\hat{X}'$ and $\hat{P}'$ can actually not commute any more here if the transformed modes 0' and 1' are to obey the standard commutation relations. In other words, the impossibility of perfect phase conjugation is reflected here by the impossibility to obtain a common eigenstate of $\hat{x}_0' - \hat{x}_1'$ and $\hat{p}_0' - \hat{p}_1'$. Instead, since $[\hat{X}', \hat{P}'] = [\hat{x}_0', \hat{p}_0'] + [\hat{x}_1', \hat{p}_1'] = 2i$, the Heisenberg uncertainty relation implies that

$$\Delta\hat{X}' \, \Delta\hat{P}' \geq \frac{1}{2} |\langle [\hat{X}', \hat{P}'] \rangle| = 1. \tag{4.41}$$

If we now assume that the phase conjugation process introduces some noise, then it is easy to determine the minimum amount of such noise for the Heisenberg uncertainty relation to be satisfied. Let us suppose that mode 1 suffers, after phase conjugation, from a random noise n_x and n_p on quadrature $\hat{x}_1'$ and $\hat{p}_1'$, respectively. Thus,

$$\hat{x}_1' = \hat{x}_1 + n_x,$$

and,

$$\hat{p}_1' = -\hat{p}_1 + n_p.$$

Naturally, we assume that this noise is unbiased, that is, $\langle n_x \rangle = \langle n_p \rangle = 0$. We also assume that the phase conjugation operation is covariant for rotations in phase-space, i. e., phase insensitive. Hence, we require the variances of n_x and n_p to be the same

($\langle n_x^2\rangle = \langle n_p^2\rangle = \sigma^2$). The resulting variance of operators $\hat{X}' = \hat{X} - n_x$ and $\hat{P}' = \hat{P} - n_p$ is

$$\Delta\hat{X}'^2 = \Delta\hat{P}'^2 = \sigma^2, \tag{4.42}$$

since $\hat{X}$ and $\hat{P}$ both have a vanishing variance in the EPR state. Equation (4.41) then implies that

$$\sigma^2 \geq 1, \tag{4.43}$$

so that the noise induced by the phase conjugation process is lower bounded by 1, i.e., *twice* the variance of a quadrature in the vacuum state ($\Delta x^2_{\rm vac} = 1/2$).

Let us now construct an actual approximate phase-conjugating transformation that attains this bound. The input mode, assumed to be prepared in a coherent state $|\alpha\rangle$, is coupled to an ancilla mode by some unitary transformation. Subsequently, the ancilla is traced over, so the processed mode is left in a mixed state that is required to be as close as possible to the complex conjugate state $|\bar{\alpha}\rangle$. Let us denote the input mode by $\hat{a}_1$ and the ancilla mode by $\hat{a}_2$. The canonical transformation can be generally described as

$$\hat{b}_i = M_{ij}\hat{a}_j + L_{ij}\hat{a}_j^*, \tag{4.44}$$

where $i, j = 1, 2$. The output modes $\hat{b}_1$ and $\hat{b}_2$ refer to the phase-conjugator output and the processed ancilla, respectively. This transformation is determined, in general, by 8 complex coefficients, but we will now impose the constraints for it to characterise an (imperfect) phase conjugator. First, we note that it is always possible to perform a phase transformation $\hat{a}_i \to e^{i\phi_i}\hat{a}_i$ and $\hat{b}_i \to e^{i\psi_i}\hat{b}_i$ such that the coefficients M_{1j} and L_{1j} are real and positive. Then, by definition, we require that the phase conjugator obeys $\langle\hat{b}_1\rangle = \langle\hat{a}_1^*\rangle$. Also, without loss of generality, we can assume that the ancilla is initially in the vacuum state $\langle\hat{a}_2\rangle = \langle(\hat{a}_2)^2\rangle = 0$ (see [46]). Thus, we must have $M_{11} = 0$ and $L_{11} = 1$. We now impose the covariance with respect to rotations, or "universality", of the transformation, that is, the constraint that the added noise is phase-insensitive (each quadrature suffers from the same noise). If the input mode has phase-insensitive noise, i.e., if $\langle(\hat{a}_1)^2\rangle = \langle\hat{a}_1\rangle^2$ (for example, if it is a coherent state), then we require that the output mode also has phase-insensitive noise, i.e., $\langle(\hat{b}_1)^2\rangle = \langle\hat{b}_1\rangle^2$. Using

$$\langle(\hat{b}_1)^2\rangle - \langle\hat{b}_1\rangle^2 = \langle(\hat{a}_1^*)^2\rangle - \langle\hat{a}_1^*\rangle^2 + M_{12}L_{12}, \tag{4.45}$$

we conclude that the universality condition (4.45) is simply $M_{12}L_{12} = 0$. Three more conditions come from imposing the commutation rules to be conserved by the transformation (4.44):

$$[b_1, b_1^*] = M_{12}^2 - L_{12}^2 - 1 = 1, \tag{4.46}$$

$$[b_2, b_2^*] = |M_{21}|^2 + |M_{22}|^2 - |L_{21}|^2 - |L_{22}|^2 = 1, \tag{4.47}$$

$$[b_1, b_2] = M_{1j}L_{2j} - L_{1j}M_{2j} = 0. \tag{4.48}$$

Equation (4.46), together with the universality condition, implies that $L_{12} = 0$ and $M_{12} = \sqrt{2}$. Equations (4.47) and (4.48) then impose two last conditions on the four coefficients M_{2j} and L_{2j}, so we are left with two free parameters. If we further impose that mode 2 transforms just as mode 1 ($M_{22} = 0$ and $L_{22} = 1$), then we get

$$\hat{b}_1 = \hat{a}_1^* + \sqrt{2}\,\hat{a}_2, \tag{4.49}$$

$$\hat{b}_2 = \sqrt{2}\,\hat{a}_1 + \hat{a}_2^*. \tag{4.50}$$

As we could expect, this transformation exactly describes a phase-insensitive phase-conjugating linear amplifier (see [46]). One can easily check that the noise variance of the output of this phase conjugator is

$$(\Delta x^2)_{b_1} = (\Delta p^2)_{b_1} = \Delta x^2_{\rm vac} + 2\Delta x^2_{\rm vac} = 3/2 \tag{4.51}$$

so that the phase-conjugation induced noise is *twice* the vacuum noise, i.e., $2\Delta x^2_{\rm vac} = 1$. Hence, this transformation is optimal as it saturates the bound (4.43). In particular, if the input is a coherent state $|\alpha\rangle$, the output will be a Gaussian mixture of coherent state ρ with variance one centred on $|\bar{\alpha}\rangle$. Consequently, the phase conjugating fidelity is

$$F = \langle\bar{\alpha}|\rho|\bar{\alpha}\rangle = 1/2, \tag{4.52}$$

just as for an optimal measurement [42, 13]. Interestingly, this implies that phase conjugation is intrinsically a classical process: it could be achieved as well by simultaneously measuring the two quadratures of $|\alpha\rangle$, and then preparing a coherent state whose quadrature p has a flipped sign. Incidentally, we note that any number of phase-conjugated outputs can actually be prepared together at no cost (with $F = 1/2$ for each).

It is interesting, at this point, to extend the parallel with the universal quantum spin-flip machine for qubits, and make a connection with a state estimation question. In Chapter 3, we have discussed the fact that encoding a direction $\mathbf{n}$ into two antiparallel spins $|\mathbf{n}, -\mathbf{n}\rangle$ yields slightly more information on $\mathbf{n}$ than encoding it into two parallel spins $|\mathbf{n}, \mathbf{n}\rangle$. Here, we investigate the counterpart of this situation for information that is carried by a continuous quantum variable instead of a qubit. Consider the situation where a sender, Alice, wants to communicate to a receiver, Bob, a complex number $\alpha = (x+ip)/\sqrt{2}$. Assume Alice is allowed to use a quantum channel only twice so as to send Bob two coherent states of a given amplitude $|\alpha|^2$ each. She can choose, for example, to send Bob the product state $|\alpha\rangle^{\otimes 2}$. In this case, the best strategy to infer both x and p with a same precision is to perform a product measurement [50] . As we have seen in Sect.(4.1), a simultaneous measurement of the two quadratures of each coherent state $|\alpha\rangle$ yields (x, p) with a variance $2\Delta x^2_{\rm vac} = 1$ [42]. The resulting error variance on x and p estimated from these two measurements is then equal to half of this variance, that is

$$\Delta x^2_{\rm vac} = 1/2.$$

Another possibility to encode α is that Alice sends Bob the product state $|\alpha\rangle \otimes |\bar{\alpha}\rangle$. In this case, a possible (but not necessarily optimal) strategy for Bob is again to carry out a product measurement, taking into account that the measured value of p of the second state should be read as $-p$. This obviously results in the same error variance $1/2$. However, the fact that the phase-conjugation transformation has a non-unity fidelity leaves open the possibility that there exists a measurement of $|\alpha\rangle \otimes |\bar{\alpha}\rangle$ that is *not* of a product form, and yields a variance strictly lower than $1/2$. Indeed, if there were a perfect universal phase conjugator, then it could be used to convert $|\bar{\alpha}\rangle$ into $|\alpha\rangle$ before applying the optimal product measurement on $|\alpha\rangle^{\otimes 2}$, thereby resulting in the same minimum variance in both cases.

Let us now explicitly describe an measurement of the product state $|\alpha\rangle \otimes |\bar{\alpha}\rangle$, which yields indeed a lower variance. Expressing the two input modes as $|\alpha\rangle = \exp(ip\hat{x}_1 - ix\hat{p}_1)|0\rangle$ and $|\bar{\alpha}\rangle = \exp(-ip\hat{x}_2 - ix\hat{p}_2)|0\rangle$, we can write the input product state as $|\alpha\rangle \otimes |\bar{\alpha}\rangle = \exp(ip\hat{X} - ix\hat{P})|0\rangle$, where $\hat{X} = \hat{x}_1 - \hat{x}_2$ and $\hat{P} = \hat{p}_1 + \hat{p}_2$ are two *commuting* operators. Assume now that the two input states $|\alpha\rangle$ and $|\bar{\alpha}\rangle$ are sent each into

one of the input ports of a balanced beam splitter, characterised by the canonical transformation

$$\hat{x}'_1 = (\hat{x}_1 + \hat{x}_2)/\sqrt{2}, \qquad \hat{p}'_1 = (\hat{p}_1 + \hat{p}_2)/\sqrt{2}, \tag{4.53}$$

$$\hat{x}'_2 = (\hat{x}_1 - \hat{x}_2)/\sqrt{2}, \qquad \hat{p}'_2 = (\hat{p}_1 - \hat{p}_2)/\sqrt{2}. \tag{4.54}$$

The input product state can be re-expressed as

$$|\alpha\rangle \otimes |\bar{\alpha}\rangle = \exp(i\sqrt{2}\, p\, \hat{x}'_2 - i\sqrt{2}\, x\, \hat{p}'_1)|0\rangle \tag{4.55}$$

implying that x and p can be measured *separately* here by applying homodyne detection on modes 1' and 2'. Indeed, a measurement of the first quadrature of mode 1' yields $\sqrt{2}\, x$, on average, while a measurement of the second quadrature of mode 2' yields $\sqrt{2}\, p$. These two measurements suffer each from an error of variance $\Delta x^2_{\rm vac} = 1/2$. Hence, the resulting error variance on x and p is reduced to

$$\Delta x^2_{\rm vac}/2 = 1/4.$$

In contrast, if we had the input product state $|\alpha\rangle^{\otimes 2}$ and were sending each coherent state $|\alpha\rangle$ into the input ports of a balanced beam splitter, we would obtain a single coherent state $|\sqrt{2}\, \alpha\rangle$ on the output mode 1'. One should then necessarily perform a *simultaneous* measurement of the two quadratures of the latter mode, yielding $(\sqrt{2}\ x, \sqrt{2}\ p)$ with an error variance $2\Delta x^2_{\rm vac} = 1$, or, equivalently x and p with a variance $\Delta x^2_{\rm vac} = 1/2$. As a consequence, we have proven here that a better strategy for sending x and p to Bob is to encode them into two conjugate coherent states $|(x + ip)/\sqrt{2}\rangle \otimes |(x - ip)/\sqrt{2}\rangle$ rather than sending two replicas of $|(x + ip)/\sqrt{2}\rangle$. The error variance on x and p is indeed reduced by a factor of two via the use of phase conjugation.

4.4 Phase-Conjugated Input Quantum cloning Machine

We now present the continuous variable analogue of the quantum cloning machine of orthogonal qubits presented in the previous chapter. We will seek for a cloning transformation that, taking as input N replicas of a coherent state $|\psi\rangle$ and N' replicas of its complex conjugate $|\bar{\psi}\rangle$, produces M optimal clones of $|\psi\rangle$. The resulting concept of phase-conjugated inputs (PCI) cloning machines will turn out to be closely connected to that of the amplification of light, just as what we found for standard cloning. As a matter of fact, a PCI cloner can be decomposed as a sequence of beam-splitters, a single non-linear medium, and another sequence of beam-splitters. We will start by deriving the optimal canonical transformation that acts on two modes in a coherent state with respective mean values $\alpha\psi$ and $\beta\bar{\psi}$ (where α, β are real while ψ is a complex number), and generates a mode whose mean value is $\gamma\psi$, where γ is real. Remarkably, this transformation will be shown to have a structure similar to that of a conventional phase-insensitive phase-preserving amplifier as defined in [46], where both the signal and idler modes are used as input. After having derived this transformation, we will apply it to the case of integer α^2, β^2, and γ^2, and see how it can be supplemented with beam-splitters to provide a PCI cloning machine for continuous variables. This machine will be shown to produce $M' = M + N' - N$ additional phase-conjugated clones (or anticlones). The quality of the clones and anticlones will be discussed in the case of a balanced cloner ($N = N'$), as well as for arbitrary phase-conjugate input fractions $N'/(N + N')$. The related question of the optimal measurement ($M = \infty$) will also be treated.

Let $\{a_i\}$ and $\{b_i\}$ $(i = 1 \ldots 3)$ denote respectively the input and output modes annihilation operators of the cloning transformation. The indices $i = 1, 2$ respectively refer to the input and phase-conjugated input modes, while $i = 3$ refers to an auxiliary mode. In full generality, we are seeking for a linear canonical transformation

$$b_i = \sum_j M_{ij} a_j + \sum_j L_{ij} a_j^* \qquad (i, j = 1 \ldots 3), \tag{4.56}$$

that meets the three following requirements. First, starting with modes a_1 and a_2 with mean values $\langle a_1 \rangle = \alpha\psi$ and $\langle a_2 \rangle = \beta\bar{\psi}$, we require $\langle b_1 \rangle = \gamma\psi$. We will only consider the case $|\gamma| \geq |\alpha|$, since, otherwise, the problem becomes trivial: one would just have to attenuate the input coherent state $|\alpha\psi\rangle$ with an unbalanced beam-splitter, yielding a coherent state of amplitude $\gamma\psi$. To simplify the problem, we may assume that $\beta = 1$, which amounts to substitute ψ for $\beta\psi$. Then, we have:

$$\begin{aligned} \alpha M_{11} + L_{12} &= \gamma, \\ M_{12} + \alpha L_{11} &= 0. \end{aligned} \tag{4.57}$$

Second, this transformation must obey the commutation rules $[b_i, b_k] = 0$ and $[b_i, b_k^*] = \delta_{ik}$ $(\hbar = 1)$, that is

$$\begin{aligned} M_{ij} L_{kj} - L_{ij} M_{kj} &= 0, \\ M_{ij} M_{kj}^* - L_{ij} L_{kj}^* &= \delta_{ik}. \end{aligned} \tag{4.58}$$

Third, the noise of the output mode b_1 of this transformation should be minimum.

Before sketching our calculation, let us note that a further simplification comes from the fact the the annihilation operators are defined up to an arbitrary phase, so that a transformation $a_i \to e^{i\mu_i} a_i$ and $b_1 \to e^{i\nu} b_1$ allows us to take M_{1j} and L_{1j} real and positive. Since we focus on phase-insensitive transformation, minimising the noise amounts to minimising the sole quantity $(\Delta b_1)^2 = \frac{1}{2}\langle b_1 b_1^* + b_1^* b_1 \rangle - \langle b_1 \rangle \langle b_1^* \rangle$ [46]. Thus, using the fact that $(\Delta a_i)^2 = 1/2$ for a mode a_i in a coherent state, we need to minimise

$$(\Delta b_1)^2 = \frac{1}{2} \sum_j (M_{1j}^2 + L_{1j}^2), \tag{4.59}$$

under the constraints Eqs. (4.57) and (4.58). Rather than solving this full problem, we use here a common trick in constrained extremisation problems that consists in solving a simpler problem with weaker constraints (bearing in mind that taking weaker constraints can only yield better solutions) and then checking that the solution of this simpler problem is one of the full problem. Specifically, we minimise $(\Delta b_1)^2$ taking into account the only condition $M_{1j} M_{1j} - L_{1j} L_{1j} = 1$. Taking Eq. (4.57) into account and introducing a Lagrange multiplier λ, we minimise the quantity $M_{11}^2 + (\gamma - \alpha M_{11})^2 + (1 + \alpha^2) L_{11}^2 + M_{13}^2 + L_{13}^2 + \lambda\big(M_{11}^2 - (\gamma - \alpha M_{11})^2 - (1 - \alpha^2) L_{11}^2 + M_{13}^2 - L_{13}^2 - 1\big)$, with respect to M_{11}, L_{11}, M_{13} and L_{13}. Some algebra shows that this problem admits only one solution $M_{13} = L_{13} = L_{11} = M_{12} = 0$, that is, the auxiliary mode is unnecessary. The optimal transformation has then the same structure as that of a phase-insensitive amplifier of gain G. Restoring β, we get

$$\begin{aligned} b_1 &= \sqrt{G} a_1 + \sqrt{G-1} a_2^*, \\ b_2 &= \sqrt{G-1} a_1^* + \sqrt{G} a_2, \end{aligned} \tag{4.60}$$

with

$$\sqrt{G} = \frac{-\alpha\gamma + \beta\sqrt{\gamma^2 - \alpha^2 + \beta^2}}{\beta^2 - \alpha^2}, \tag{4.61}$$

It can be easily checked that, for $\beta = 0$ (or $\alpha = 0$), Eq. (4.60) reduces to a phase-insensitive phase-preserving (or phase-conjugating) amplifier as defined in [46], and can be used to carry out the $N \to M$ cloning (or phase-conjugating) transformation described in the previous section.

Let us now turn to the special case where α^2, β^2 and γ^2 are integers (which we will denote respectively as N, N', M). The transformation Eq. (4.60) can be used as the central element of a PCI cloning machine, which is covariant for translations and rotations in phase space (see Fig. 4.4). Indeed, the following procedure can be used to produce M optimal clones of a coherent state $|\psi\rangle$ from $|\psi\rangle^{\otimes N}|\bar{\psi}\rangle^{\otimes N'}$:

(i) Concentrate the N replicas of $|\psi\rangle$ stored in the N modes $\{c_l\}$ $(l = 0 \dots N-1)$ into a single mode a_1, this results in a coherent state of amplitude $\sqrt{N}\,\psi$. This operation can be performed with a network of beam-splitters achieving a N-mode Discrete Fourier Transform (DFT), which yields the mode

$$a_1 = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} c_l, \tag{4.62}$$

and $N-1$ vacuum modes. Similarly, concentrate the N' replicas of $|\bar{\psi}\rangle$ stored in the N' modes $\{d_l\}$ $(l = 0 \dots N'-1)$, into a single mode a_2 in a coherent state of amplitude $\sqrt{N'}\,\bar{\psi}$, with the help of an N'-mode DFT:

$$a_2 = \frac{1}{\sqrt{N}} \sum_{l=0}^{N'-1} d_l. \tag{4.63}$$

(ii) Process the modes a_1 and a_2 into a "phase-conjugated inputs" amplifier (PCIA), resulting in modes b_1 and b_2 as defined in Eqs. (4.60) and (4.61).

(iii) Distribute the output b_1 into M clones $\{c'_l\}$ $(l = 0 \dots M-1)$ with a M-mode DFT:

$$c'_l = \frac{1}{\sqrt{M}}(b_1 + \sum_k e^{i\pi kl/M} v_k), \tag{4.64}$$

where $\{v_k\}$ $(k = 1 \dots M-1)$ denote $M-1$ vacuum modes. It is readily verified that this procedure yields M clones of $|\psi\rangle$.

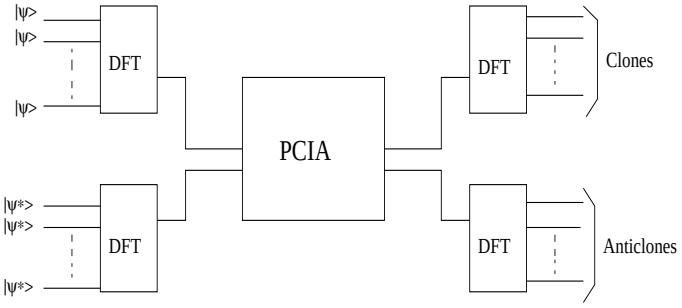

Figure 4.3: PCI cloner that produces M clones and M' anticlones from N replicas of $|\psi\rangle$ and N' replicas of $|\bar{\psi}\rangle$. The modes are concentrated and distributed by a Discrete Fourier Transform (DFT) and amplified in a phase-conjugate input modes amplifier (PCIA).

Interestingly, the amplitude b_2 of the other output of the PCIA has a mean value $\sqrt{M'}\bar{\psi}$, with

$$N - N' = M - M'. \tag{4.65}$$

Therefore, it can be used to produce M' phase-conjugated clones (or anticlones) of $|\psi\rangle$, $\{d_l'\}$ $(l = 0 \dots M' - 1)$, using a M'-mode DFT:

$$d_l' = \frac{1}{\sqrt{M'}}(b_2 + \sum_k e^{i\pi kl/M} w_k) \tag{4.66}$$

where $\{w_k\}$ $(k = 1 \dots M' - 1)$ denote $M' - 1$ vacuum modes. Clearly, this procedure is optimal to produce M clones since its central element, the PCIA, is optimal, and the beam-splitters are passive elements. In addition, the M' anticlones that are produced at no cost are also optimal. Indeed, our transformation produces M optimal clones and $M \geq N$, and is symmetric with respect to the interchange of labels 1 and 2. So, if our initial problem was to produce M' optimal anticlones with $M' \geq N'$, we would find the same solution. Since $M \geq N \iff M' \geq N'$, it is clear that our transformation yields both optimal clones and optimal anticlones. Furthermore, since the PCIA is linear and phase-insensitive, the resulting PCI cloner is covariant with respect to translations and rotations of the state to be copied: all coherent states are copied equally well, and the cloning-induced noise is the same for all quadrature components.

Using Eqs. (4.60)-(4.64) and (4.66), the noise of the clones and anticlones can be written as

$$(\Delta c_l')^2 = \frac{1}{2} + \frac{G-1}{M}, \quad (\Delta d_l')^2 = \frac{1}{2} + \frac{G-1}{M'}, \tag{4.67}$$

where the gain can be re-expressed as a function of the number of inputs and outputs,

$$\sqrt{G} = \frac{\sqrt{N'M'} - \sqrt{NM}}{N' - N} \tag{4.68}$$

As expected, the variance of the output clones exceeds 1/2, implying that the clones are not exactly in the coherent state $|\psi\rangle$. Instead, their state is given by

$$\frac{1}{\pi\sigma_c^2} \int d^2\beta e^{-|\beta|^2/\sigma_c^2} D(\beta)|\psi\rangle\langle\psi|D(\beta)^*, \tag{4.69}$$

where $\sigma_c^2 = (G-1)/M$.

Consider now the balanced case ($N = N'$, $M = M'$), for which simple analytical expressions of the noise variances can be obtained. Taking the limit $\alpha \to \beta$ in Eq. (4.61) and replacing α^2 by N and γ^2 by M yields $G = (M+N)^2/4MN$, so that the error variances of the clones and anticlones are

$$(\Delta c_l')^2 = (\Delta d_l')^2 = \frac{1}{2} + \frac{(M-N)^2}{4M^2N}. \tag{4.70}$$

Note that this balanced cloner is optimal amongst all PCI cloners in the sense that it minimises σ_c^2 for fixed $N+N'$ and $M+M'$. It is convenient to characterise the quality of cloning in terms of the fidelity $f_{\binom{N}{N}\to M} = \langle\psi|\rho_c|\psi\rangle/|\langle\psi|\psi\rangle|^2$ where ρ_c denotes the state of the clones. Using Eq. (4.69), we get

$$f_{\binom{N}{N}\to M} = \frac{1}{1+\sigma_c^2} = \frac{4M^2N}{4M^2N + (M-N)^2}. \tag{4.71}$$

Let us now compare the production of M clones from N replicas and N antireplicas to the production of M clones from $2N$ identical replicas. The variance and fidelity of the clones k_i obtained by standard cloning are given by

$$(\Delta k_i')^2 = \frac{1}{2} + \left(\frac{1}{2N} - \frac{1}{M}\right), \tag{4.72}$$

and

$$f_{2N\to M} = \frac{2MN}{2MN + M - 2N}. \tag{4.73}$$

Of course, in the trivial case where $M = 2N$, standard cloning can be achieved perfectly, while the balanced PCI cloner yields an additional variance $\sigma_c^2 = 1/(16N)$. However, whenever $M \geq 2N + 1$, the $\binom{N}{N} \to M$ balanced cloner always yields a lower variance (hence a higher fidelity) than the $2N \to M$ cloning machine. The balanced PCI cloner is also better for the anticlones: more anticlones are produced at no cost, and they have a better fidelity. Indeed, a standard $2N \to M$ cloning machine produces $M - 2N$ anticlones of fidelity $2N/(2N+1)$, which actually is the fidelity of an optimal measurement of $2N$ replicas of $|\psi\rangle$. In contrast, a PCI cloner produces M anticlones with a higher fidelity, as given by Eq. (4.71). In particular, for $M \to \infty$, we see from Eqs. (4.70) and (4.72) that the additional noise induced by a PCI cloner is $1/4N$, that is, half the noise induced by a standard $2N \to \infty$ cloner $(1/2N)$. Note that in this case, the output of the PCIA can be considered as classical and the underlying process appears to be equivalent to a measurement. This reflects that *more classical information can be encoded in N pairs of phase-conjugated replicas of a coherent state than in 2N identical replicas*, as discussed in the previous section. More generally, in the unbalanced case ($N \neq N'$), it can be shown that the optimal measurement results in a noise that is equal to that obtained by measuring $(\sqrt{N} + \sqrt{N'})^2$ identical replicas of the input, in the absence of phase-conjugated inputs.

We have shown that the balanced PCI cloner can result in better cloning quality than a standard cloner. More generally, we may ask the following question: *If we want to produce M clones of a coherent state $|\psi\rangle$ from a fixed total number n of input modes, N of which being in the coherent state $|\psi\rangle$ and N′ of which being in the phase-conjugated state $|\bar{\psi}\rangle$, what is the phase-conjugate fraction $a = N'/n$ that minimises the error variances of the clones?*

From Eq. (4.60), we see that σ_c^2 then depends only on a, and varies as

$$G(a) = \left(\frac{\sqrt{a}\sqrt{\frac{M}{n} + (2a-1)} - \sqrt{\frac{M}{n}}\sqrt{1-a}}{2a-1} \right)^2 \tag{4.74}$$

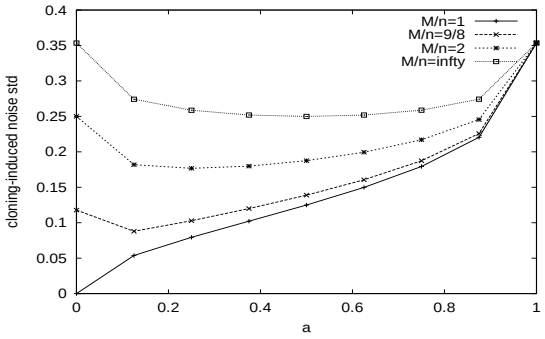

Figure 4.4: Cloning-induced noise standard deviation $\sqrt{\sigma_c^2}$ as a function of the phase-conjugate fraction $a = N'/n$, for $n = 8$ and several values of M/n.

In Fig. (4.4), we have plotted $\sqrt{\sigma_c^2}$ as a function of a for $n = 8$ and different values of $M \geq n$. In the trivial case where $M = n = 8$, the minimum additional variance is

of course zero, and is obtained for $a = 0$. The cloning transformation is then just the identity. However, when $M \geq n+1$, using phase-conjugated input modes yields lower variances than standard cloning if a is correctly chosen (the lowest variance is attained for $a \neq 0$). Remarkably, the value of a achieving the minimum variance is not equal to $1/2$ for finite M, that is the optimal input partition contains more replicas than antireplicas. In the limit of large M, however, the number of antireplicas achieving the lowest variances tends to $n/2$, and the curve $G(a)$ tends to a symmetric curve around $a = 1/2$. This symmetry is not surprising, since $M = \infty$ corresponds to a measurement and we expect that measuring the value of ψ from N replicas of $|\psi\rangle$ and N' replicas of $|\bar{\psi}\rangle$ is equivalent to starting from N' replicas of $|\psi\rangle$ and N replicas of $|\bar{\psi}\rangle$. Finally, in the case where $a = 1$, the transformation consists in producing M clones of $|\psi\rangle$ from n replicas of $|\bar{\psi}\rangle$. This is just phase-conjugation. The additional variance is therefore given by $1/n$, which does not depend on M. This explains why the curves converge all to the same point at $a = 1$.

4.5 Summary

We have studied the issue of quantum cloning for continuous variables. We have considered the case where one wants to clone all coherent states equally well. Our figure of merit for cloning was the added noise in x quadrature variance and the added noise in p quadrature variance. We have seen that optimal quantum cloning can then be achieved with Gaussian operations such as amplification and mixing of modes with beam splitters. We have essentially considered symmetric cloners. This is the first step in understanding how information contained in quantum systems distributes. An interesting extension of our work would be the study of asymmetric machines for which the fidelities of the clones are not equal. Such a study should allow to understand further how information contained in one (or several) quantum system(s) distributes in more quantum systems.

Interestingly, the cloners we have derived are not optimal for all figures of merit. While writing up this thesis, Patrick Navez and Nicolas Cerf have discovered that if one considers the fidelity as a figure of merit instead of the added noise, there exists a non-Gaussian cloner outperforming the transformations presented here. Considering the case of duplication, there is a cloner achieving a fidelity of 0.6825 instead of 2/3 in our case.

A cloning transformation using phase conjugate input modes has also been considered. Again, this transformation has been shown to be decomposable in a sequence of beam splitters, a central amplification stage, and another sequence of beam-splitters. A possible way to implement this central stage would be to use four-wave mixing. Two weak fields entering the $\chi^{(3)}$ medium would then play the role of the phase-conjugated inputs, and energy would be brought to the system by two external modes in a large coherent state [51]. We have shown evidence that PCI cloning transformations outperform standard cloning transformations (taking only identical inputs) if the goal is to produce clones and anticlones of a state or to get knowledge about a state through measurement. A possible extension of our work would be to study the case where the number of anticlones is a free parameter (in the PCI cloner derived here, it is constrained by N, N', and M). An interesting generalisation would be asymmetric PCI cloning transformations. We could then determine whether there is still an advantage in having conjugate input modes rather than identical ones.

Chapter 5

Quantum cryptography I: Protocols

We review the main principles of Quantum Key Distribution. We start by describing some of the motivations of quantum cryptography and the BB84 protocol. Then we describe the squeezed-state protocol and the coherent-state protocol to which the next chapter is devoted.

5.1 Cryptography

Cryptography aims at providing two parties, an emitter and a sender, with secure means of confidential communication. Cryptographic techniques are numerous, but the general scheme is always the same. The message to be transmitted is first encoded with a key by the emitter. The message is sent through some channel at the end of which lies the authorised receiver. Finally, the receiver decodes the message with a key, which may be either identical or different from the one used for encoding. Cryptographic protocols divide into two classes: the public key protocols and the secret key protocols [9]. Being very practical, public key protocols are the most often encountered in everyday life. They are for example widely used on the Internet. Unfortunately, they suffer from a major weakness: they rely on computability assumptions. These assumptions are highly plausible but unproven. RSA [5], for example, the most widely used public key protocol, would be seriously endangered if an algorithm existed to decompose efficiently an integer into its prime factors. Although, the best known algorithms run on classical computers for factorisation require computation times which are (sub-)exponential in the size of the number to factorise, as indicated by Eq. (1.1), there is no guarantee that no efficient classical algorithm exists for factorisation.

Vernam Encryption

Most secret key protocols suffer from the same flaws. The only protocol we are perfectly sure of is the Vernam encryption, or "one-time pad", which works as follows. Suppose the (clear) message to be communicated from Alice to Bob[1] is a finite sequence of bits $\mathcal{M}$. Vernam Encryption is operated with a key $\mathcal{K}$, that is another sequence of bits, which is as long as $\mathcal{M}$.

[1] Traditionally in cryptography, the authorised emitter and sender are respectively called "Alice" and "Bob", and a potential eavesdropper is referred to as "Eve".

- Alice encrypts the message by adding it to the key (modulo 2):

$$\mathcal{E} = \mathcal{M} \oplus \mathcal{K}.$$

- $\mathcal{E}$ is sent to Bob through a classical channel.
- Bob decrypts the message:

$$\mathcal{M} = \mathcal{E} \oplus \mathcal{K}.$$

This protocol is very simple indeed. Its security is guaranteed by the fact that if the key $\mathcal{K}$ is fully random, then the encrypted message $\mathcal{E}$ is fully random too, whatever the message $\mathcal{M}$ is. Thus, $\mathcal{E}$ brings no information about $\mathcal{M}$ to a potential eavesdropper, Eve, if $\mathcal{K}$ is private. It only remains to find efficient ways to distribute the key. Quantum Cryptography or Quantum Key Distribution addresses precisely this issue.

5.2 The BB84 protocol

Quantum cryptography fundamentally relies on a beautiful idea. The use of non-orthogonal quantum states to carry the key bits makes it possible to detect any potential eavesdropping. The resources needed for QKD always comprise a source of non-orthogonal quantum states on Alice's side, a quantum channel conveying these states to Bob, a measuring apparatus on Bob's side, and a (public) authenticated classical channel between Alice and Bob, see Fig.5.2. QKD protocols generally consists in two (intertwined) parts. The first part consists of probing the quantum channel to determine whether it is possible to securely transmit the key over it. The second part consists of the explicit distillation of the secret key.

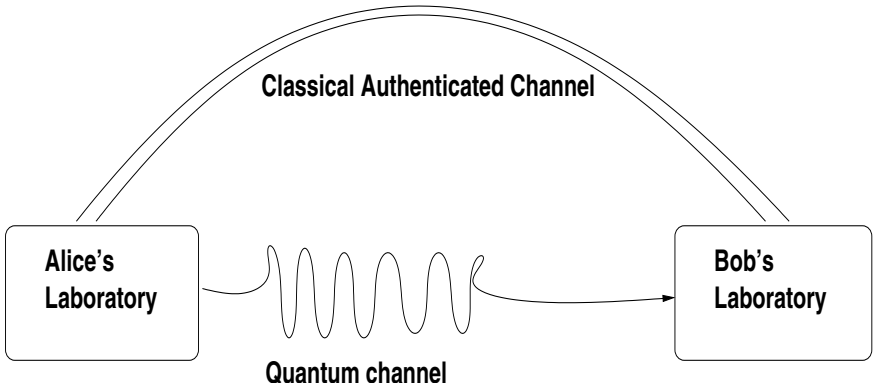

Figure 5.1: Overall cryptographic scheme.

The first QKD protocol, BB84, was invented by Bennett and Brassard in 1984 [52]. This protocol is at the root of all the remainder of this chapter and the next chapter.

The protocol

Let $\mathcal{H}$ denote the Hilbert space of a qubit. The BB84 protocol makes use of two orthonormal bases of $\mathcal{H}$, b_0 and b_1, to transmit a key. $\mathsf{b}_0 \equiv \{|e_{00}\rangle, |e_{01}\rangle\}$ and $\mathsf{b}_1 \equiv$

$\{|e_{10}\rangle, |e_{11}\rangle\}$ are defined such that

$$|\langle e_{bs}|e_{b's'}\rangle| = (1 - \delta_{bb'})/\sqrt{2} + \delta_{bb'}\delta ss'. \tag{5.1}$$

This condition expresses the fact that the bases b_0 and b_1 are mutually unbiased. That is, if a state of one basis is measured in the other basis, the outcome will be fully random. Stated otherwise, a measurement in the basis b_1 of a state $|e_{0k}\rangle$ of b_0 doesn't bring any information about $|e_{0k}\rangle$ and vice versa.

If, for example, the qubit is encoded in the polarisation state of a single photon, b_0 could represent a set of two orthogonal directions of polarisation, which we decide to call "horizontal" and "vertical". And b_1 would then represent another set of two orthogonal polarisation states, mutually unbiased with the states of the first set, say "circular" and "anticircular".

The BB84 protocol is run as follows:

- 1. Alice sends (about) $4N$ qubits to Bob through the quantum channel. The ith qubit, $i = 1 \ldots 4N$, is prepared in the state $|e_{b(i)s(i)}\rangle$, where the values $b(i)$ and $s(i)$ are drawn randomly and independently from a uniform distribution.

- 2. $\forall i = 1 \ldots 4N$, Bob draws a random bit $b'(i)$, and measures the ith qubit sent by Alice in the basis $\mathsf{b}_{b'(i)}$. Let $\{s'(i)\}$ denote the sequence of outcomes he obtains.

- 3. Through the classical authenticated (public) channel, Alice reveals in which basis each qubit has been prepared; $\{b(i)\}_{i=1\ldots4N}$.

- 4. Alice and Bob discard the qubits for which Bob's choice of basis for measurement doesn't match Alice's choice of basis for preparation ($b(i) \neq b'(i)$). The number of such qubits should be (about) $2N$.

- 5. On a subset T of the remaining qubits ($|T| \approx N$), Alice and Bob probe the quantum channel. If the quantum channel is noiseless, Alice's preparation and Bob's outcome will agree for all qubits in T. If the quantum channel is noisy, possibly due to the intervention of an eavesdropper, Alice and Bob will find a non-zero error-rate δ_b.

- 6. Classical post-processing. If δ_b is small enough, (i) Alice and Bob apply some classical communication protocol to correct errors and obtain a sifted key K_{sif}, and (ii) Alice and Bob apply some classical communication protocol to extract a private key from K_{sif}.

The figure (5.2) illustrates (a part of) the BB84 protocol.

Eavesdropping

We have deliberately defined the last step of the protocol loosely. Actually, the security of the BB84 protocol crucially depends on the classical post-processing. This dependence is quite complicated and we here only want to describe QKD protocols. Classical post-processing will be addressed in greater detail in the next chapter, at the same time as general security to which it is intimately related.

Nonetheless, the essential features of the BB84 protocol that make it robust can already be illustrated in the case of a simple eavesdropping strategy. First note that it is only after Bob has received all the qubits that (i) Alice reveals in which basis each qubit has been prepared (ii) Alice and Bob agree on which qubits will serve to probe

State prepared by Alice	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\circlearrowleft$	$\leftrightarrow$	$\circlearrowright$	$\updownarrow$	$\circlearrowleft$
Alice's raw key	0	0	1	0	0	0	1	0
Basis chosen by Bob	+	◯	+	◯	+	+	+	+
Bob's outcome	$\leftrightarrow$	×	$\updownarrow$	$\circlearrowleft$	$\leftrightarrow$	×	$\updownarrow$	×
Bob's sifted key	0	-	1	0	0	-	1	-

Figure 5.2: Illustration of the BB84 protocol. Qubits are encoded by polarisation states of single photons, and sent via a noiseless quantum channel. + means that the photon is measured in the horizontal-vertical basis. ◯ means that the photon is measured in the circular-anticircular basis. $\{\leftrightarrow, \circlearrowleft\} \rightarrow$ "0", $\{\updownarrow, \circlearrowright\} \rightarrow$ "1", "$4N$" $= 8$.

the quantum channel. Whatever strategy the eavesdropper, Eve, uses, it never brings her any advantage not treating all qubits equally.

Let us assume now that Eve, adopts the "intercept and resend strategy" to get information about the raw key. She intercepts each qubit. She decides randomly to measure it either in the b_0 basis or in the b_1 basis. She prepares a new qubit, according to her result, and sends it to Bob. This strategy brings her as much information as Bob about the raw key, but it doesn't leave her undetected.

Note that only the qubits for which Alice's and Bob's choice of basis agree should be examined. The other qubits are discarded anyway. For each remaining qubit, two scenarios are possible. First scenario: Alice's, Bob's and Eve's choices of bases agree. Eve knows the bit and is left undetected. Second scenario: Alice's and Bob's choices of bases agree, but disagree with Eve's. Let us illustrate this situation with an example. Let us assume that Alice prepares a qubit in the state $|e_{00}\rangle$ and that Eve measures it in the basis b_1. With a probability equal to $1/2$, she will get the result $'0'$ and so send Bob the state $|e_{10}\rangle$. When measuring this state in the basis b_0, Bob will get the result '1' with a probability equal to $1/2$, and infer that the state prepared by Alice was $|e_{01}\rangle \neq |e_{00}\rangle$. A similar situation occurs if Alice sends the state $|e_{01}\rangle$ (resp. $|e_{10}\rangle$ or $|e_{11}\rangle$), with Bob measuring in the b_0 basis (resp. b_1) and Eve measuring in the b_1 basis (resp. b_0). So, the protocol is robust against the intercept and resend strategy since Eve will tag (about) a quarter of the eavesdropped qubits. An error rate δ exceeding (about) 25% informs Alice and Bob that the quantum channel is noisy and might be tapped by an eavesdropper. They can then abort the protocol. Fig.5.2 illustrates the situation.

State prepared by Alice	$\leftrightarrow$	$\leftrightarrow$	$\updownarrow$	$\circlearrowleft$	$\leftrightarrow$	$\circlearrowright$	$\updownarrow$	$\circlearrowleft$
Alice's raw key	0	0	1	0	0	1	1	0
Basis chosen by Eve	◯	+	+	+	◯	◯	+	◯
Basis chosen by Bob	+	◯	+	◯	+	+	+	+
Bob's outcome	$\updownarrow$	$\circlearrowleft$	$\updownarrow$	$\circlearrowright$	$\leftrightarrow$	$\updownarrow$	$\updownarrow$	$\leftrightarrow$
Bob's sifted key	!1!	-	1	!1!	0	-	1	-

Figure 5.3: Illustration of the BB84 protocol in the case where an eavesdropper adopts an intercept and resend strategy. About a quarter of the qubits are tagged.

Let us now assume that Eve runs a weaker attack. Let us assume that she only

taps the quantum channel and runs her intercept-and-resend strategy with a non-unit probability p. In such a case, δ will be about $0.25p$ but she will only know (about) $p|K_{raw}|$ bits of the raw key though Bob's knowledge of the raw key will be $|K_{raw}|$ bits. This advantage of $(1-p)|K_{raw}|$ bits could then be exploited by Alice and Bob to extract a secret private key.

Can we conclude that an error rate below 25% means that the quantum channel is safe? No. Eve can actually adopt strategies more efficient than intercept and resend. A first better strategy for Eve is to use (an asymmetric variant of) the quantum cloning machine presented in Chapter 3 to try copying instead of measuring the states sent by Alice. This brings the security threshold down to 15%. Is there yet any better strategy? In fact, we don't know. However, as we shall see in the next chapter, if the error rate is below 11%, the protocol can be made secure by suitable classical post-processing.

5.3 Quantum Key Distribution with Continuous Variables

Squeezed-state protocols

The BB84 protocol has been originally proposed for qubits, but it is easy to generalise it to higher-dimensional systems. We could for example let the key elements be carried by harmonic oscillators, $\mathcal{H} = L^2(\mathbf{R})$. These harmonic oscillators could be physically represented by a single mode of a quantised electromagnetic field. In principle, we could modify the BB84 protocol to a Quantum Key Distribution (QKD) protocol where, instead of bits, the key elements are real values drawn from a "uniform distribution" over $\mathbf{R}$, and where the pair of bases are $\hat{x}$ quadrature eigenstates $|x\rangle$ and $\hat{p}$ quadrature eigenstates $|p\rangle$. This modified protocol would retain all the ingredients of the BB84 protocol, and hence work as well.

However, $\hat{x}$ eigenstates and $\hat{p}$ eigenstates are unphysical nonnormalisable states. Neither is it possible to draw the key elements from a uniform probability distribution over $\mathbf{R}$. A regular version of the BB84 protocol with oscillators was proposed by Cerf et al. [53, 54] and reads as follows.

- Alice sends Bob (about) $4N$ quantum oscillators, each prepared in a squeezed state. To prepare the ith squeezed state, $i = 1 \dots 4N$, Alice draws a random bit $b(i)$ to determine whether the state is squeezed in $\hat{x}$ (or in $\hat{p}$), as well as a Gaussian-distributed random variable X_A of variance Σ_x^2 (or P_A of variance Σ_p^2) to determine the centre of the squeezed state. According to her result, x (or p), she sends an x-squeezed state centred on $(x, 0)$ (or a p-squeezed state centred on $(0, p)$).

- Bob receives the $4N$ squeezed states. For each of them, he draws a random bit to determine whether he measures the $\hat{x}$ or the $\hat{p}$ quadrature, and performs the measurement. Let his result be denoted by a random variable X_B.

- Alice and Bob discard the oscillators for which Bob's measurement doesn't match Alice's preparation. The number of such oscillators should be (about) $2N$.

- On a subset T of the remaining oscillators ($|T| \approx N$), Alice and Bob probe the quantum channel relating them. If the quantum channel is too noisy, they abort the protocol.

- Classical post-processing. Alice and Bob apply some classical communication protocol to extract (i) a sifted key K_{sif} from their correlated real values (ii) a private key from K_{sif}.

Like in BB84, half the measurements give results that are uncorrelated to Alice's values, so half of the samples must be discarded when Alice discloses the encoding variable. Unlike BB84, however, measuring the correct variable does not yield the exact value of the key element, r, even with a perfect apparatus, because of the intrinsic noise of the squeezed state. The value r follows a Gaussian distribution $N(0, \Sigma_{x,p})$, to which some Gaussian noise is added $N(0, \sigma_{x,p})$, thus resulting in a Gaussian distribution with variance $\Sigma^2_{x,p} + \sigma^2_{x,p}$. We can therefore model the transmission of r as a classical Gaussian channel with a signal-to-noise ratio (SNR) equal to Σ^2_x/σ^2_x or Σ^2_p/σ^2_p.

An important requirement of the protocol is to make it impossible for Eve to be able to infer which encoding variable Alice used. For this, measuring the correct or incorrect variable (x or p) must yield statistically indistinguishable results. If, in contrast, Eve was able to detect (even not perfectly) whether she measured the wrong variable, then she could get and exploit this information to improve her attack. This indistinguishability requirement can be expressed as the equality of the density matrices resulting from the two encoding rules. This requirement reads

$$\sigma^2_x + \Sigma^2_x = \frac{1}{4\sigma^2_p},$$
$$\sigma^2_p + \Sigma^2_p = \frac{1}{4\sigma^2_x}. \tag{5.2}$$
$$\tag{5.3}$$

This also means that the SNR is the same for both variables x and p, and that the maximal information rate is given by Shannon's formula for the classical capacity of a Gaussian channel with constrained input power Σ^2/σ^2 [55]:

$$I = \frac{1}{2}\log_2(1 + \Sigma^2_x/\sigma^2_x) = -\log_2(2\sigma_x\sigma_p). \tag{5.4}$$

This information is non-zero provided that the $x-$ or $p-$states (or both) are squeezed below the shot noise limit ($\sigma^2_{x,p} < 1/2$).

Eavesdropping by cloning. Let us now discuss an individual eavesdropping of this protocol[2] with cloning machines such as those defined by Eq.(4.12). Eve makes two clones of the state sent by Alice, one of which is transmitted to Bob, and the other is measured in the correct variable when Alice reveals the encoding rule. In fact, this happens to be the optimal individual eavesdropping strategy as shown in [53] and [56].

We use a $1 \to 2$ cloning machine, and we keep the freedom to make a better clone for Bob or Eve (parameter χ) and to get more accuracy in x or p (parameter λ). The subscripts 1 and 2 for the two copies are replaced respectively by B and E for the two recipients. According to Eq. (4.12), the added variances on the clones will be:

$$\begin{aligned} \sigma^2_{B,x} &= \frac{1}{2}\chi\lambda,\ \sigma^2_{B,p} = \frac{1}{2}\chi\lambda^{-1}, \\ \sigma^2_{E,x} &= \frac{1}{2}\chi^{-1}\lambda,\ \sigma^2_{E,p} = \frac{1}{2}\chi^{-1}\lambda^{-1}. \end{aligned} \tag{5.5}$$

Let us calculate the resulting information rates. When Bob measures x, the result is affected both by the intrinsic fluctuations of x and by the noise induced by the cloning

[2]Individual eavesdropping means that Eve probes the key elements one by one. More general strategies can be considered.

operation, thus resulting in a total variance $\sigma_x^2 + \frac{1}{2}\chi\lambda$. This is the noise variance in the Gaussian channel representing the communication between Alice and Bob through Eve's cloning machine. Therefore, the information rate is now

$$I_{B,x} = \frac{1}{2}\log_2(1 + \frac{\Sigma_x^2}{\sigma_x^2 + \frac{1}{2}\chi\lambda}). \tag{5.6}$$

Similarly, one can calculate the new variance of p measured by Eve on her clone, namely $\sigma_p^2 + \frac{1}{2}\chi^{-1}\lambda^{-1}$. This gives an information rate

$$I_{E,p} = \frac{1}{2}\log_2(1 + \frac{\Sigma_p^2}{\sigma_p^2 + \frac{1}{2}\chi^{-1}\lambda^{-1}}). \tag{5.7}$$

Adding the last two information rates indicates the balance between Bob's and Eve's information. Remarkably, the information that Eve gains by using this attack on p is exactly equal to the information that Bob loses on x [53],

$$I_{B,x} + I_{E,p} = \frac{1}{2}\log_2(1 + \frac{\Sigma_x^2}{\sigma_x^2}) = I. \tag{5.8}$$

Of course, this balance also works when swapping x and p, namely $I_{B,p} + I_{E,x} = I$.

This (fairly intuitive) result is interesting because it allows Bob to upper bound the information gained by a possible eavesdropper. Assuming symmetry of the protocol in x and p, Bob can estimate $I - I_B$ and is then guaranteed that $I_E \leq I - I_B$. One can prove [57] that with reconciliation and privacy amplification carried out over a public authenticated channel, one is guaranteed to generate secret key bits whenever

$$I_B - I_E > 0. \tag{5.9}$$

This last condition is in turn guaranteed provided that $I_B > I/2$, so that up to a 50% information loss on Bob's side is acceptable in order to generate key bits. In particular, an eavesdropping with $\chi \geq 1$ generates at least 50% of information loss so that it makes the scheme insecure.

Coherent-state protocols

As we have shown, the construction of squeezed states protocols follow, in a sense, from the requirement of having a continuous variable protocol mimicking the original BB84 protocol [53, 58]. Though these protocols are physically sensible, they rely on the preparation of squeezed states, which is experimentally quite a demanding task. It is thus desirable to extend the squeezed-state protocol to a protocol using only *coherent states*. To meet this requirement, Grosshans and Grangier proposed a protocol [59], the security of which we will study in the next chapter. The trick with this protocol, which makes it secure, is that now Alice modulates *both* the $\hat{x}$ and the $\hat{p}$ mean values of the state she sends. The protocol runs as follows:

- Alice sends Bob (about) $2N$ quantum oscillators, each prepared in a coherent state. To prepare the ith coherent state, $i = 1 \ldots 2N$, Alice draws two random real numbers x_A and p_A from a Gaussian distributed law with variance Σ^2. The values x_A and p_A determine the centre of the ith coherent state sent by Alice.
- Bob receives the $2N$ coherent states. For each of them, he randomly decides to measure either $\hat{x}$ or $\hat{p}$, and performs the measurement.
- Bob declares publicly which quadrature he measured for each oscillator.

- On a subset T of the sent oscillators, Alice and Bob probe the quantum channel relating them.
- Classical post-processing. If the channel relating Alice and Bob is judged safe, they can extract (i) a sifted key K_{sif} from their correlated real values (ii) a private key from K_{sif}.

Eavesdropping. A first security analysis can be carried using the same cloning machines, as for the squeezed-state protocol.In contrast to Eq.(5.4), no squeezing is necessary now. Because of the modulation of both quadratures for each key element, we are no longer constrained by Eq.(5.2).

The information sent by Alice is still given by Shannon formula. So,

$$I = \frac{1}{2}\log(1 + \frac{\Sigma^2}{1/2}). \tag{5.10}$$

Assuming that Eve treats both quadratures equally, a calculation similar to that we have just carried for squeezed states shows that [59]:

$$I_B - I_E = \frac{1}{4}\log\frac{(\Sigma^2 + 1/2 + \chi)}{\chi} - \frac{1}{4}\log\frac{(\Sigma^2 + 1/2 + 1/\chi)}{1/\chi}, \tag{5.11}$$

which, just as for the squeezed state protocol, is nonzero as long as $\chi < 1$.

5.4 Summary

Quantum cryptography allows two parties to get a secret key. Ideally, it should rely on no assumption about the resources of a potential eavesdropper. Focusing on squeezed state protocols and coherent state protocols, we have seen that these protocols are robust against individual attacks.

From an experimental point of view, qubit based QKD protocols such as BB84 have proven their reliability over quite long distances, typically of order of 10 km, when implemented using (approximations of) single photons source[3]. But the bit rates achieved by such protocols are deceptively low. Typically, a secret key generation rate of 1000 bit/sec is possible over a distance of 70km (optical fibre implementation, 1550 nm, loss $\approx 0.5dB/km$) [9]. The essential reasons for such low rates are the necessity to *simulate* single photon sources at Alice's side, and the imperfection of detectors at Bob's side [9]. However, a recent experiment, using genuine single photon sources, gives hope for more efficient implementations of BB84 [60].

On the other hand, continuous variable protocols, and in particular coherent state protocols, such as those presented here seem to be a promising alternative to qubit based protocols. They seem to allow for facilitated implementations, over larger distances and higher key-generation rates (see [10] and references therein). It is therefore crucial to study their security. This is the aim of the next chapter.

[3]To build a single photon source is experimentally quite a challenging task.

Chapter 6

Quantum Cryptography II: Security analysis

The security proof of the BB84 quantum key distribution protocol against collective attacks is reviewed, as well as shift-resistant quantum codes, and the construction that connects the latter with the former to devise secure squeezed-state protocols. We then analyse the security of coherent-state protocols.

6.1 Collective Attacks

In the preceding chapter, we have given a hint of why quantum cryptography is secure. For example, we have shown the robustness of the coherent-state protocol against a specific class of attacks. However, a quantum cryptographic scheme should ideally be founded only on physical assumptions and, unlike classical cryptography, should rely on no assumption about the resources of a potential eavesdropper. But is it possible?

Let us look at the global scheme of a quantum cryptographic setup, see Fig.5.2 . This setup is made of four parts: Alice's laboratory, Bob's laboratory, the quantum channel and the classical channel. Our first observation is that if the classical channel relating Alice and Bob is not indeed authenticated, no quantum key distribution is possible. Nothing then stops Eve from pretending to be Bob with respect to Alice and Alice with respect to Bob. Therefore, we do make the first assumption:

Assumption 6.1.1 *The classical channel between Alice and Bob is authenticated, i.e. the classical messages sent from Alice to Bob arrive unaffected and vice versa.*

Second, we further limit Eve's power as follows:

Assumption 6.1.2 *Eve has no control over Alice's laboratory and Bob's laboratory.*

We make this assumption because we find it physically sensible and because again nothing is anymore possible if Alice's and Bob's laboratories are under Eve's control. We note nevertheless that it is sometimes useful to partially relax Assumption (6.1.2). For example, one can then prove that the BB84 protocol is still robust when Alice and Bob don't have perfect experimental apparati (see [61] and references therein).

Assumptions 6.1.1 and 6.1.2 were implicitly made when we presented cryptographic protocols in the previous chapter. What distinguishes the analysis that shall be presented here from that of the previous chapter is that we now want to make as few hypotheses as possible about the way Eve controls the quantum channel. In particular, we want our security analysis to still hold if Eve runs a collective attack, i.e. an

attack where she might eavesdrop several key elements at a time. It is tempting to conjecture that Alice drawing the key elements values and encoding choice independently from identical probability distributions should imply that no collective attack is more informative to Eve than an individual attack. However, such a conjecture has, to date, never been proven for any QKD scheme. Another distinction from the previous chapter, is that we would like our analysis to apply (or at least to be extendible) to the case of non-Gaussian attacks (Gaussian attacks are those mapping a Gaussian state to a Gaussian state). Finally, another interest of the security analysis carried out here is that it will lead us to be more precise about the classical post processing part of a QKD protocol, on which the security crucially relies, as we shall see.

The analysis of security under individual attacks has been made possible by an explicit construction of optimal eavesdropping strategy. Such a construction is highly likely to be untractable for collective attacks. The approach that we will adopt now is different. The key idea underlying the security analysis that we will carry now is the following trivial fact, which is a direct consequence of Schmidt decomposition (see chapter 2):

> Let $|\Psi\rangle_{ABE}$ denote a tripartite pure quantum state. If the subsystem $\rho_{AB} = \mathrm{Tr}_E |\Psi\rangle_{ABE}\langle\Psi|$ is pure, then $|\Psi\rangle_{ABE}$ is of the form $|\phi\rangle_{AB} \otimes |\mu\rangle_E$. That is, E is factored out and can thus get no information about the shared randomness that A and B will extract from $|\phi\rangle_{AB}$.

6.2 The Shor-Preskill proof for BB84

The Shor-Preskill (SP) proof of security of the BB84 protocol will be our starting point. This proof contains two ingredients:

- Establishing the security of an entanglement purification based protocol (this term will be explained below).
- Establishing the equivalence between this entanglement purification based protocol and the BB84 protocol.

Thus, one concludes that the BB84 protocol *augmented with suitable classical post-processing* is secure.

CSS codes

The entanglement purification protocols of interest for us are based on the use of a particular class of quantum codes: the CSS codes, after Calderbank-Shor-Steane [62, 63]. As we will see, these codes have the nice feature of decoupling the correction of bit-errors and the correction of phase-errors. This feature will turn particularly useful when turning an entanglement-purification protocol into the BB84 protocol.

Let $\mathcal{H}^{\otimes l}$ denote the Hilbert space of l qubits ($\mathcal{H} = \mathbf{C}^2$). A CSS code Q is defined from two embedded classical linear codes C_1 and C_2:

$$\{0\} \subset C_2 \subset C_1 \subset \mathbf{F}_2^l, \tag{6.1}$$

$$\{0\} \subset C_1^\perp \subset C_2^\perp \subset \mathbf{F}_2^l. \tag{6.2}$$

Let l_1 (resp. l_2) denote the dimension of C_1 (resp. C_2). Q is the linear span of vectors

$$|v + C_2\rangle \equiv \frac{1}{|C_2|^{1/2}} \sum_{w \in C_2} |v + w\rangle, \; v \in C_1. \tag{6.3}$$

The encoded information is "which coset of C_2 in C_1", so that the dimension of the protected subspace is $\dim Q = 2^{\dim C_1 - \dim C_2}$. Let us introduce the tensor product operators

$$X_a^s = X_a^{s_1} \otimes \ldots \otimes X_a^{s_l}, \tag{6.4}$$

and

$$Z_a^s = Z_a^{s_1} \otimes \ldots \otimes Z_a^{s_l}, \tag{6.5}$$

where $s = (s_1, \ldots, s_l) \in \mathbf{F}_2^l$. Let H_1 denote the parity check matrix of C_1 and $H_2^\perp$ denote the parity check matrix of $C_2^\perp$.

Let us first consider bit-flip errors and states of the form $|v\rangle = |v_1\rangle \ldots |v_l\rangle$, with $v \in C_1$. Bit-flips can be represented as

$$E_\epsilon^b : |v\rangle \to |v + \epsilon\rangle, \tag{6.6}$$

where E_ϵ^b is of the form X_a^s. To recover from bit-flips, we apply a joint operation on the affected state and an ancilla:

$$|v + \epsilon\rangle \otimes |0\rangle_{anc} \to |v + \epsilon\rangle \otimes |H_1(v + \epsilon)\rangle_{anc} = |v + \epsilon\rangle \otimes |H_1\epsilon\rangle_{anc}. \tag{6.7}$$

Measuring the ancilla then gives the bit-error syndrome. This measurement corresponds to a measurement of operators of the form X^{s_i}, where $\{s_i\}$ are the rows of H_1. We can then proceed as for classical linear codes: we recover by applying an X operator on the qubits the syndrome has identified as affected. Thus, if the code C_1 has distance $d \geq 2t_b + 1$, the code corrects up to t_b bit-flip errors. Finally, since codewords (6.3) are linear combinations of such states $|v\rangle$, applying the transformation (6.7) on a codeword of Q will indeed allow to diagnose the bit-flip error (and later correct it) without damaging quantum information.

To describe the correction of phase-flip errors, it is useful to describe how a codeword transforms under the application of the l-tensor Hadamard transformation $H^{\otimes l}$, where

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{6.8}$$

Direct computation shows that a state $|w\rangle = |w_1\rangle \ldots |w_l\rangle$, $w \in \mathbf{F}_2^l$, transforms under $H^{\otimes n}$ as $|w\rangle \to \sum_{u \in \mathbf{F}_2^l} (-)^{u \cdot w} |u\rangle$. Hence, a codeword $|w + C_2\rangle$ transforms according to

$$|w + C_2\rangle \to \frac{1}{\sqrt{2^{n - l_2}}} \sum_{u \in C_2^\perp} (-)^{u \cdot w} |u\rangle, \tag{6.9}$$

where we have used the identity

$$\sum_{v \in C_2} (-)^{u \cdot v} = \begin{cases} 0 & \text{if } u \in C_2^\perp, \\ 2^{l_2} & \text{if } u \notin C_2^\perp. \end{cases} \tag{6.10}$$

Phase-flip errors can be represented as

$$E_\epsilon^p : |v\rangle \to (-)^{v \cdot \epsilon} |v\rangle, \tag{6.11}$$

where E_ϵ^p is of the form Z_a^s. For recovery, we first apply $H^{\otimes l}$, which transforms an affected codeword $E_\epsilon^p|w+C_2\rangle$ as

$$E_\epsilon^p|w+C_2\rangle \rightarrow H^{\otimes l}E_\epsilon^p|w+C_2\rangle = \sum_{u\in C_2^\perp}(-)^{u\cdot w}|u+\epsilon\rangle, \tag{6.12}$$

where we have used Eqs (6.9)(6.10). Phase errors are corrected exactly as bit errors but syndromes are now diagnosed with respect to the code $C_2^\perp$. The diagnostic reads

$$|u+\epsilon\rangle\otimes|0\rangle_{anc} \rightarrow |u+\epsilon\rangle\otimes|H_2^\perp\epsilon\rangle_{anc}. \tag{6.13}$$

Measuring the ancilla then gives the phase-error syndrome. This measurement corresponds to a measurement of operators of the form Z^{s_i}, where $\{s_i\}$ are the rows of $H_2^\perp$. Once the affected qubits have been identified by the syndrome, two equivalent procedures are possible for recovery: (i) Apply X on the affected qubits and apply the (inverse) Hadamard transformation to get the recovered state. (ii) First apply H, and then apply Z on the affected qubits. If the code $C_2^\perp$ has distance $d_2^\perp \geq 2t_p+1$, we can recover up to t_p phase errors. Finally, Y errors are just combinations of X and Z errors. So the distance of a CSS code satisfies

$$d \geq min(d_1, d_2^\perp). \tag{6.14}$$

Let us now explain why we said that CSS codes decouple bit error correction and phase error correction. The relations (6.1)(6.2) imply that $H_2^\perp H_1^t = 0$. Thus, for all operator X^{s_1}, $s_1 \in H_1$, and for all Z^{s_2}, $s_2 \in H_2^\perp$, measured to diagnose bit flip errors (resp. phase flip errors), we have:

$$[X^{s_1}, Z^{s_2}] = 0, \tag{6.15}$$

i.e. *correcting bit errors doesn't disturb phases and vice versa.*

What is the *rate* of such codes? If $e_b = t_b/l$ and $e_p = t_p/l$ respectively denote the bit-error rate and the phase-error rate of the quantum channel, one can prove that there exist CSS codes that will be (asymptotically) successful in reliably transmitting quantum information if:

$$R \equiv 1 - h(e_b) - h(e_p) > 0, \tag{6.16}$$

where $h(x) = -\log_2 x^x(1-x)^{(1-x)}$ denotes the binary Shannon entropy [21]. When $e_b = e_p$, this rate hits zero for $e_b = 11\%$.

Finally, we can, in analogy to what has been done in the classical case, define a "translated" code $Q_{x,z}$ of a CSS code Q. $Q_{x,z}$ is the linear span of vectors of the form:

$$|v+C_2\rangle_{x,z} \equiv \frac{1}{|C_2|^{1/2}}\sum_{w\in C_2}(-)^{z\cdot w}|x+v+w\rangle,\ v\in C_1. \tag{6.17}$$

Example. A simple example of a CSS code can be constructed from the 7-bit Hamming code presented in chapter 2. It is the 7-qubit code introduced by Steane [63]. Let us remind that the Hamming 7-bit code C is a linear code whose parity check matrix is given by

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}, \tag{6.18}$$

and a generator matrix for this code is

$$G = \begin{pmatrix} H \\ (v) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.19}$$

We have $C^\perp \subset C$ and C decomposes into the two cosets of $C^\perp$ in C: $C = \{C^\perp\} \cup \{v + C^\perp\}$. We can thus construct a CSS quantum code from the Hamming code if we set $C_1 = C$ and $C_2^\perp = C_1$, implying $C_2^\perp = C$. Therefore, the Hamming parity check matrix H can be used to diagnose both bit flips (when working in the computational basis) and phase flips (when working in the Hadamard basis). A basis of this code is given by

$$|\bar{0}\rangle_F = \frac{1}{\sqrt{8}} \sum_{w \in C^\perp} |w\rangle, \tag{6.20}$$

$$|\bar{1}\rangle_F = \frac{1}{\sqrt{8}} \sum_{w \in v + C^\perp} |w\rangle. \tag{6.21}$$

Since both $|0\rangle_F$ and $|1\rangle_F$ are superpositions of codewords of C, bit flips can be diagnosed with an H parity check. Also, since

$$|\bar{0}\rangle_P = H^{\otimes 7}|\bar{0}\rangle_F = \frac{1}{\sqrt{2}}(|\bar{0}\rangle_F + |\bar{1}\rangle_F) = \frac{1}{4}\sum_{v \in C}(-)^{0 \cdot v}|v\rangle \tag{6.22}$$

$$|\bar{1}\rangle_P = H^{\otimes 7}|\bar{0}\rangle_F = \frac{1}{\sqrt{2}}(|\bar{0}\rangle_F - |\bar{1}\rangle_F) = \frac{1}{4}\sum_{v \in C}(-)^{1^7 \cdot v}|v\rangle, \tag{6.23}$$

and since the relation (6.12) holds, we see that a phase flip can be diagnosed as a bit flip in the dual basis with the parity check matrix H (Note that Eq.(6.23) holds because $(1111111) \in v + C^\perp$).

The seven qubit code is robust against any single bit and phase error on anyone of the seven qubits. But error recovery will fail if two different qubits both undergo a bit flip error or a phase flip error. Consider for example the case where two qubits undergo a bit flip error. Let e_1, e_2 denote two different weight-one strings. $He_1 + He_2$ is the sum of two columns of H, and therefore another column of H. Therefore, there exists a third weight-1 vector e_3 such that$H(e_1 + e_2 + e_3) = 0$. $e_1 + e_2 + e_3$ is weight-3 (because at most weight-3 as a sum of three weight-1 vectors, and at least weight-3 as a codeword of H). A corrector might diagnose e_3 as the actual error, and effect e_3, resulting in an overall weight-3 error $e_1 + e_2 + e_3$, thus inducing a bit flip.

Entanglement Purification

An entanglement purification protocol is a procedure by which two remote parties (Alice and Bob) want to extract k' pure EPR pairs:

$$|\phi_0\rangle^{\otimes l'} \equiv (\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle))^{\otimes l'}$$

from a state close to l perfect EPR pairs ($l \geq l'$), using only local operations and classical communication.

It is natural that quantum error-correcting codes yield entanglement purification procedures. After all, a state close to l noisy EPR pairs can be seen as l perfect EPR pairs whose halves have undergone some noise. Let us make this statement more precise and construct entanglement purification protocols using CSS codes.

First suppose that Alice and Bob start with l perfect EPR pairs, all in the state $|\phi_0\rangle$. Let them both measure Z^r, for each row $r \in H_1$. From Alice and Bob's point of view, the measured syndromes (resp. s_z^A and s_z^B) are completely random but the *relative* syndrome $s_z^A - s_z^B$ always equals zero. Similarly, when they measure X^r, for each row $r \in H_2$, Alice and Bob get random but correlated results $s_x^A = s_x^B$. Now if x and z denote any l-bit string vectors such that $H_1 x = s_z^A = s_z^B$ and $H_2^\perp z = s_x^A = s_x^B$, we see that what Alice and Bob have done when measuring all the syndromes is projecting their state $|\phi_0\rangle^{\otimes l}$ onto the state $|\phi_0\rangle^{\otimes l'}$ encoded by $Q_{x,z}$ ($l' \leq l$). Now what if Alice and Bob do not start with l perfect EPR pairs? Let us define

$$|\phi_1\rangle = (\mathbf{1} \otimes X)|\phi_0\rangle, \tag{6.24}$$

$$|\phi_2\rangle = (\mathbf{1} \otimes Z)|\phi_0\rangle, \tag{6.25}$$

$$|\phi_3\rangle = (\mathbf{1} \otimes XZ)|\phi_0\rangle. \tag{6.26}$$

Let t_b and t_p denote respectively the number of bit errors and the number of phase errors that the codes $Q_{x,z}$ can correct. Suppose that Alice and Bob share a state with t_b or fewer bit flips ($|\phi_2\rangle$ or $|\phi_3\rangle$) and t_p or fewer phase flips ($|\phi_1\rangle$ or $|\phi_3\rangle$). Then Alice and Bob can purify their noisy entangled pairs by measuring their bit (phase) error syndrome, computing their relative bit (phase) error syndrome and correcting the corresponding bit (phase) errors to get the state $|\phi_0\rangle^{\otimes l'}$ encoded by $Q_{x,z}$.

QKD with EPR pairs

Let us now show how this entanglement purification (EP) protocol can be extended to a secure QKD protocol. Let us describe the (EP+QKD) protocol as a whole.

Protocol 1: QKD based on entanglement purification #1 Alice creates the state $|\phi_0\rangle^{\otimes 2l}$ #2 Alice selects a random $2l$-bit string b and performs a Hadamard transformation on the second half of each EPR pair for which $b = 1$. #3 Alice sends Bob the second half of each of the $2l$ pairs. #4 Bob acknowledges receipt of his $2l$ halves. #5 Alice selects randomly l check pairs to test Eve's interference. #6 Alice reveals b and which pairs are check pairs. #7 Bob performs a Hadamard transformation on the qubits for which $b = 1$. #8 Alice and Bob measure their halves of the check pairs in the Z basis, and compare their results. If too many outcomes mismatch, they abort the protocol. #9 Alice and Bob apply the above-described entanglement purification protocol to the remaining code qubits and get l' (nearly) perfect pairs $|\phi_0\rangle^{\otimes l'}$. #10 Alice and Bob measure these perfect EPR pairs in the Z basis and get an l'-bit secret key.

To prove that this protocol is secure, we will need the following theorem.

Theorem 6.2.1 (Lo and Chau) *[64] If Alice and Bob share a state $\rho \in \mathcal{B}(\mathcal{H})$ such that the fidelity of ρ with m pure EPR pairs $|\phi_0\rangle^{\otimes m}$ satisfy*

$$\mathrm{tr}(\rho(|\phi_0\rangle\langle\phi_0|)^{\otimes m}) \geq 1 - 2^{-s}, \tag{6.27}$$

then Eve's mutual information with the key is at most $2^{-c} + 2^{0(-2s)}$, where $c = s - \log_2(2m + s + 1/\ln 2)$.

Qualitatively, the more the state shared by Alice and Bob is pure, the smaller the information Eve gets about the key.

In Protocol 1, Alice and Bob estimate e_b and e_p from the check qubits. If e_b and e_p satisfy Eq.(6.16), they judge that the EP protocol is likely to work, and operate further. How reliable is it to proceed this way? Let us calculate the probability that the test on the check qubits succeeds while the EP on the code qubits fails. First of all, since we have assumed that Eve doesn't control Alice's laboratory, she has no means to know which qubits are check qubits and which are code qubits before Bob receives them. It is therefore legitimate to assume that she treats all qubits equally. According to Eq.(2.18), this means that Eve's intervention on each pair can be modelled as

$$|\phi_0\rangle_{AB}|0\rangle_E \to |\phi_0\rangle|e_{\mathbf{1}}\rangle + |\phi_1\rangle|e_X\rangle + |\phi_2\rangle|e_Z\rangle + |\phi_3\rangle|e_Y\rangle, \tag{6.28}$$

whatever the precise nature of this intervention is, that is whatever the states

$$|e_{\mathbf{1}}\rangle, |e_X\rangle, |e_Z\rangle, |e_Y\rangle$$

are. Let ρ_{AB}^{ind} denote the individual state of each pair Alice and Bob share. One could thus conclude that Alice and Bob should perform Bell measurements[1] on each check pairs to estimate

$$e_b = \text{tr}(\rho_{AB}^{\text{ind}}(|\phi_1\rangle\langle\phi_1| + |\phi_3\rangle\langle\phi_3|)) \tag{6.29}$$
$$e_p = \text{tr}(\rho_{AB}^{\text{ind}}(|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|)). \tag{6.30}$$

But fortunately, the identities

$$|\phi_1\rangle\langle\phi_1| + |\phi_3\rangle\langle\phi_3| = |01\rangle\langle 01| + |10\rangle\langle 10|, \tag{6.31}$$
$$|\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3| = |+-\rangle\langle +-| + |-+\rangle\langle -+|, \tag{6.32}$$

where $|\pm\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$, show that the estimations of e_b and e_p will be exactly the same if Alice and Bob perform only local measurements in the X and Z basis.

We now show that the EP protocol applied to the l pairs produces a state that is close to l' encoded EPR pairs $|\overline{\phi_0^{\otimes l'}}\rangle$. Let us decompose $\mathcal{H}^{\otimes l}$, the Hilbert space of l qubits into

$$\mathcal{H}^{\otimes l} = \mathcal{H}_g \oplus \mathcal{H}_g^{\perp},$$

where $\mathcal{H}_g$ is the linear span of vectors of the form $\otimes_{i=1}^{l}(\mathbf{1}\otimes X^{a_i})\otimes_{i=1}^{l}(\mathbf{1}\otimes Z^{b_i})|\phi_0\rangle^{\otimes l}$, with $a_i, b_j = 0, 1$ and $\sum_i a_i \leq t_b, \sum_j b_j \leq t_p$. $\mathcal{H}_g$ is the linear span of EPR pairs having undergone at most t_b bit-flips, and at most t_p phase-flips. Let Π denote the projector onto $\mathcal{H}_g$ and $\Pi^{\perp}$ denote the projector onto $\mathcal{H}_g^{\perp}$, and let $\rho \in \mathcal{B}(\mathcal{H}^{\otimes l})$ denote the state shared by Alice and Bob. $\text{tr}\,\rho\Pi$ (resp. $\text{tr}\,\rho\Pi^{\perp}$) is the probability that ρ has support on $\mathcal{H}_g$ (resp. $\mathcal{H}_g^{\perp}$). Let ρ' denote the state obtained after error correction. If ρ were projected onto $\mathcal{H}_g$, then the fidelity of ρ' with $|\overline{\phi_0^{\otimes l'}}\rangle$ is 1, and if ρ were projected onto $\mathcal{H}_g^{\perp}$, this fidelity would be some number rate f_π , $0 \leq f_\pi \leq 1$. Thus on average, the fidelity of ρ' with $|\overline{\phi_0^{\otimes l'}}\rangle$ will be

$$\langle\overline{\phi_0^{\otimes l'}}|\rho'|\overline{\phi_0^{\otimes l'}}\rangle = \text{tr}\,\rho\Pi + f_\pi\,\text{tr}\,\rho\Pi^{\perp} \geq \text{tr}\,\rho\Pi. \tag{6.33}$$

As we have said, Eve has no means to treat check qubits differently than code qubits. The check qubits are therefore a faithful sample of code qubits. We can thus invoke classical probability theory which tells us that the probability to have δl bit (phase) errors on the code qubits and fewer than $(\delta - \epsilon)l$ errors on the check bits is asymptotically less than $\exp\left(-\epsilon^2 l/4(\delta - \delta^2)\right)$. Let us pause and summarise.

[1] A Bell measurement is a measurement on the Hilbert space of two qubits whose resolution of unity reads $|\phi_0\rangle\langle\phi_0| + |\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2| + |\phi_3\rangle\langle\phi_3|$.

(i) The probability that Alice and Bob pass the verification test, and that the state they get after error-correction is not exponentially close to $|\phi_0\rangle^{l'}$ is exponentially small. (ii) If the fidelity of the state Alice and Bob share at the end of the EP protocol is exponentially close to $|\phi_0\rangle^{l'}$, then Eve's mutual information with the key is exponentially small. (i) & (ii) If Alice and Bob pass the verification test, then Eve's mutual information with the key is exponentially small: the EP based QKD protocol is secure.

Equivalence with the BB84 protocol

We now see how Protocol 1 can be converted to the BB84 protocol without loss of security. This conversion proceeds in two steps. Protocol 1 is first converted to a protocol where we still use CSS codes but where we don't need entangled pairs anymore (Protocol 2). Then we give the arguments turning this protocol into the BB84 protocol.

EP-based protocol to CSS-based protocol. In protocol 1, Alice and Bob use two types of pairs: check pairs and code pairs. For both types, it doesn't matter whether Alice performs her measurement on her half before or after transmitting his half to Bob. If she measures the check qubits before transmission, this is the same as preparing the state $|0\rangle$ or $|1\rangle$ and sending it to Bob. Also, if she measures the syndromes on the code qubits before transmission, this is the same as choosing two random vectors $x, z \in \mathbf{F}_2^l$, a random key $k \in \mathbf{F}_2^l$, and to send the state $|k\rangle$ encoded by the (translated) CSS code $Q_{x,z}$. The EP-based protocol is thus equivalent to the following protocol.

Protocol 2: CSS-based protocol. #1 Alice creates l random check qubits, a random l'-bit key k, and a random string $b \in \mathbf{F}_2^{2l}$. #2 Alice draws two vectors $x, z \in \mathbf{F}_2^l$ randomly, according to a uniform distribution. #3 Alice encodes the key k with the CSS code $Q_{x,z}$. #4 Alice chooses l positions (out of $2l$) where she puts the check bits. She puts the code bits in the remaining positions. #5 Alice applies a Hadamard transformation on each qubit i for which $b_i = 1$. #6 Alice sends the resulting state to Bob, who acknowledges receipt of the qubits. #7 Alice reveals b, the positions of the check bits, and the vectors x and z. #8 Bob undoes the Hadamard transformation in each qubit i for which $b_i = 1$. #9 Bob checks whether too many of the check bits have been damaged, and aborts the protocol if so. #10 Bob can decode the qubits and use them for the key.

CSS-based protocol to BB84. The structure of CSS codes will now be advantageously exploited to turn Protocol 2 into BB84. After Alice and Bob are confident that e_b and e_p satisfy Eq. (6.16), Bob will apply H on the code qubits for which $b = 1$ and obtain a state from which he will extract the key. But to do so, he doesn't need to correct the phase errors of this state. Let $k' \in C_1$ denote some vector such that $k' + C_2 = k$. The encoded state sent by Alice to Bob reads:

$$|k' + C_2\rangle_{x,z} \equiv \frac{1}{|C_2|^{1/2}} \sum_{\alpha \in C_2} (-)^{z\cdot\alpha} |k' + \alpha + x\rangle, \; v \in C_1. \tag{6.34}$$

Now, if Bob doesn't need to correct phase errors, we can assume that Alice doesn't send z to Bob. The (mixed) state that Bob then actually sees is the state (6.34) averaged over z, that is

$$\begin{aligned} \Xi(k' + C_2) &= \frac{1}{2^n |C_2|} \sum_z \left[\sum_{\alpha,\beta \in C_2} (-)^{(\alpha+\beta)\cdot z} |k' + \alpha + x\rangle\langle k' + \beta + x| \right] rate \\ &= \frac{1}{|C_2|} \sum_{\alpha \in C_2} |k' + \alpha + x\rangle\langle k' + \alpha + x|, \end{aligned} \tag{6.35}$$

where we have used again Eq.(6.10).

To summarise, Alice sends Bob $|k'+\alpha+x\rangle$ over the quantum channel, with $k'+\alpha \in C_1$ and $x \in \mathbf{F}_2^l$. Bob performs a measurement on the state he receives and gets the result $k'+\alpha+x+\epsilon$. Alice reveals the correction information x. Bob calculates $k'+\alpha+x+\epsilon+x = k'+\alpha+\epsilon$ and corrects the result to a codeword of C_1, which is $k'+\alpha$ with high probability (error correction or sifting). The key is $k'+\alpha+C_2 = k'+C_2$ (privacy amplification).

To get the BB84 protocol, let us make a further remark. In protocol 2, Bob stores his qubits in a quantum memory and waits until Alice reveals b to perform a Hadamard transformation on each qubit i for which $b_i = 1$, and make a measurement in the Z basis. Alternatively, he could measure these qubits in the X basis. But thanks to Eq.(6.35), the state Bob sees before Alice reveals b now reads

$$\otimes_{i=1}^{l} H^{b_i} \Xi(k' + C_2) \otimes_{i=1}^{l} H^{b_i}. \tag{6.36}$$

This state is a tensor product of pure qubit states, each either eigenstate of X, or eigenstate of Z. Consequently, Bob doesn't need any quantum memory and the protocol works as well if Alice sends Bob twice more qubits, Bob decides randomly to measure each either in the X basis, or in the Z basis and Alice and Bob discard each qubit for which Bob's choice of measurement doesn't match Alice's choice of preparation. Now setting $k'+\alpha+x = v$, v is random word of $\mathbf{F}_2^l$, and $x = u+v$ is another word of $\mathbf{F}_2^l$ such that $u \in C_1$, we get the BB84 protocol augmented with suitable classical post-processing.

Protocol 3: BB84. #1 Alice creates a random $4(l+\delta)$-bit string. #2 She chooses a random $4(l+\delta)$-bit string b. For each bit, she creates a random Z-eigenstate if b indicates 0, and a random X-eigenstate if b indicates 1. #3 Alice sends the resulting qubits to Bob. #4 Bob receives the $4(l+\delta)$ qubits and measures each in the Z basis or in the X basis at random. #5 Alice reveals b. #6 Bob discards the qubits for which his choice of basis doesn't match that of Alice. With high probability, they are left with (about) $2l$ qubits. Alice decides at random on l check bits for verification. #7 Alice and Bob announce the values of their check bits and estimate e_b and e_p. If these two quantities don't satisfy Eq. (6.16), they abort the protocol. #8 Alice announces $u+v$ where v is the vector formed by the remaining code bits, and u is a random codeword of C_1. #9 From his code bits vector, $v+\epsilon$, Bob calculates $v+\epsilon+u+v = u+\epsilon$ and corrects to the nearest codeword of C_1 (error correction). #10 The key is the coset $u+C_2$ (privacy amplification).

6.3 Shift-resistant codes

We now introduce shift-resistant codes [65]. We will later need these codes to present the construction of squeezed-state protocols secure against collective attacks [58].

Shift-resistant codes aim at reliably encode qubits using continuous variable systems. As hinted by their names, these codes are typically robust against translations in phase-space. Even if the Hilbert space embedding the code is no more a tensor product of qubits Hilbert spaces, we can invoke the formalism of stabiliser codes to describe shift-resistant codes. A shift resistant code $\mathcal{R}(0,0)$ is the simultaneous eigenspace of the two (commuting) operators

$$S_x = e^{i2\sqrt{\pi}x}, S_p = e^{-i2\sqrt{\pi}p}, \tag{6.37}$$

with eigenvalue(s) $S_x = S_p = 1$. Thus, the allowed values of x and p in the code $\mathcal{R}(0,0)$ are integer multiples of $\sqrt{\pi}$, and the codewords are invariant under shifts in x

or p by integer multiples of $2\sqrt{\pi}$. An orthogonal basis for the logical qubits is

$$\begin{aligned} |\bar{0}\rangle &= \sum_{s=-\infty}^{\infty} |x = (2s)\sqrt{\pi}\rangle = \sum_{s=-\infty}^{\infty} |p = s\sqrt{\pi}\rangle, \\ |\bar{1}\rangle &= \sum_{s=-\infty}^{\infty} |x = (2s+1)\sqrt{\pi}\rangle = \sum_{s=-\infty}^{\infty} (-)^s |p = s\sqrt{\pi}\rangle. \end{aligned} \tag{6.38}$$

The logical operators

$$\bar{Z} = e^{i\sqrt{\pi}x}, \bar{X} = e^{-i\sqrt{\pi}p} \tag{6.39}$$

commute with the stabiliser generators and act on the codewords as

$$\bar{Z}|\bar{0}\rangle = |\bar{0}\rangle \quad ; \quad \bar{Z}|\bar{1}\rangle = -|\bar{1}\rangle, \tag{6.40}$$

$$\bar{X}|\bar{0}\rangle = |\bar{1}\rangle \quad ; \quad \bar{X}|\bar{1}\rangle = |\bar{0}\rangle. \tag{6.41}$$

$\mathcal{R}(0,0)$ is robust against x-shifts Δx and p-shifts Δp that satisfy

$$|\Delta x| < \sqrt{\pi}/2, \ |\Delta p| < \sqrt{\pi}/2. \tag{6.42}$$

Errors are diagnosed by measuring the stabiliser generators (6.37). When the values of x and p are determined modulo $\sqrt{\pi}$, a displacement is applied in phase space to adjust x and (resp. p) to the nearest integer multiple of $\sqrt{\pi}$. The condition 6.42 can be simply interpreted: the error zones around the peaks of $|\bar{0}\rangle$ and $|\bar{1}\rangle$ should not overlap.

As usual, we can define (equivalent) translated codes. We will denote them $\mathcal{R}(\phi_x, \phi_p)$. They are associated with the eigenvalues of the stabiliser generators $S_x = e^{2\pi i\phi_x}$ and $S_p = e^{-2\pi i\phi_p}$. In this code, logical operators read

$$\bar{Z}(\phi_x) = e^{i\sqrt{\pi}(x-\phi_x\sqrt{\pi})}, \ \bar{X}(\phi_p) = e^{-i\sqrt{\pi}(p-\phi_p\sqrt{\pi})}. \tag{6.43}$$

A basis of $\mathcal{R}(\phi_x, \phi_p) : \{|\bar{0}(\phi_x, \phi_p)\rangle, |\bar{1}(\phi_x, \phi_p)\rangle\}$ can be obtained by applying the translation operator $e^{i\sqrt{\pi}\hat{x}\phi_p} e^{-i\sqrt{\pi}\hat{p}\phi_x}$ to the basis $\{|\bar{0}\rangle, |\bar{1}\rangle\}$ of the code $\mathcal{R}$. We have

$$|\bar{0}(\phi_x, \phi_p)\rangle = \sum_s e^{i\pi(2s+\phi_x)\phi_p} |x = (2s+\phi_x)\sqrt{\pi}\rangle \tag{6.44}$$

$$|\bar{1}(\phi_x, \phi_p)\rangle = \sum_s e^{i\pi(2s+1+\phi_x)\phi_p} |x = (2s+1+\phi_x)\sqrt{\pi}\rangle \tag{6.45}$$

Finally, we can define asymmetric shift resistant codes by stabiliser generators

$$S_x = e^{i2\sqrt{\pi}x/\alpha}, S_p = e^{-i2\sqrt{\pi}p\alpha}. \tag{6.46}$$

Such codes are robust against errors

$$|\Delta x| < \alpha\sqrt{\pi}/2, \ |\Delta p| < \sqrt{\pi}/2\alpha. \tag{6.47}$$

The codewords (6.38) are nonnormalisable states. One way to regularise them is to replace them by the Gaussian approximations

$$|\tilde{0}\rangle \approx (\frac{4}{\pi})^{1/4} \int dx|x\rangle e^{-1/2\Delta_p^2 x^2} \sum_s e^{-(x-2s\sqrt{\pi})^2/2\Delta_x^2} \tag{6.48}$$

$$\approx (\frac{1}{\pi})^{1/4} \int dp|p\rangle e^{-1/2\Delta_x^2 p^2} \sum_s e^{-(p-s\sqrt{\pi})^2/2\Delta_p^2}, \tag{6.49}$$

$$|\tilde{1}\rangle \approx \bar{X}|\tilde{0}\rangle. \tag{6.50}$$

If Δ_x and Δ_p are small, then $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$ will be close to the ideal codewords $|\bar{0}\rangle$ and $|\bar{1}\rangle$. The states $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$ can be seen as ideal codewords having been processed through a quantum channel $\mathcal{G}_{\Delta_x,\Delta_p}$ effecting

$$\begin{aligned} \mathcal{G}_{\Delta_x,\Delta_p}: \quad & |\bar{0}\rangle\langle\bar{0}| \to |\tilde{0}\rangle\langle\tilde{0}|, \\ & |\bar{1}\rangle|\bar{1}\rangle \to |\tilde{1}\rangle\langle\tilde{1}|. \end{aligned} \tag{6.51}$$

Taking $\Delta_x = \Delta_p = \Delta$ for example, the probability P_e that $\mathcal{G}_{\Delta,\Delta}$ effects an uncorrectable error, that is, the probability that $\mathcal{G}_{\Delta,\Delta}$ has shifted $|\bar{0}\rangle\langle\bar{0}|$ into the error zones of $|\bar{1}\rangle\langle\bar{1}|$, is bounded by the probability that $\mathcal{G}_{\Delta,\Delta}$ has effected a shift in x exceeding $\sqrt{\pi}/2$, i.e.

$$P_e \leq \frac{2}{\sqrt{\pi\Delta^2}} \int_{\sqrt{\pi}/2}^{\infty} dx\, e^{-x^2/\Delta^2} \leq \frac{2\Delta}{\pi} e^{-\pi/4\Delta^2}. \tag{6.52}$$

For an asymmetric code, taking $\Delta_q = \Delta\alpha$ and $\Delta_p = \Delta/\alpha$, the probability of uncorrectable q-errors is bounded by

$$P_e^x \leq \frac{2\Delta\alpha}{\pi} e^{-\pi/4(\Delta\alpha)^2}, \tag{6.53}$$

and the probability of uncorrectable p-errors is bounded by

$$P_e^p \leq \frac{2\Delta}{\alpha\pi} e^{-\pi\alpha^2/4\Delta^2}. \tag{6.54}$$

6.4 A secure squeezed-state protocol

The BB84 protocol makes use of qubits, but doesn't depend on the manner these qubits are represented. In particular, Alice and Bob could encode qubits in continuous variable systems using the shift-resistant codes we have just described. Thus the following protocol is a secure implementation of BB84.

Protocol 4: BB84 with shift-resistant codes. #1 Alice creates a $4(l+\delta)$-bit string. #2 She chooses a random $4(l+\delta)$-bit string b. For each bit, she draws randomly two numbers $\phi_x, \phi_p \in [-1/2; 1/2[$ from a uniform distribution , and creates a random Z eigenstate of the code $\mathcal{R}(\phi_x, \phi_p)$ if b indicates 0, and a random X eigenstate of the code $\mathcal{R}(\phi_x, \phi_p)$ if b indicates 1. #3 Alice sends the resulting qubits to Bob. #4 Bob acknowledges receipt. #5 Alice reveals which code $\mathcal{R}(\phi_x, \phi_p)$ she used for each qubit. #6 Bob measures each qubit either in the Z basis, either in the X basis. #7 Alice reveals b. Next steps are identical to Protocol 3.

Protocol 4 requires sophisticated manipulations from Alice and Bob: Alice should prepare complicated encoded states, Bob should store each state sent by Alice in a quantum memory until she reveals the values ϕ_x, ϕ_p, and Bob should make measurements of $\hat{x} \bmod \sqrt{\pi}$ and $\hat{p} \bmod \sqrt{\pi}$. We now show that this protocol can be simplified without loss of security. Suppose Bob chooses to measure in the Z basis, say. First, he doesn't need any quantum memory, neither does he need to measure the operators (6.43). He can as well measure the observable $\hat{x}$ on the state he receives, store the classical outcome x in a classical memory, subtract $\phi_x\sqrt{\pi}$ from x when Alice reveals ϕ_x, and adjust $x - \phi_x\sqrt{\pi}$ to the nearest integer multiple of $\sqrt{\pi}$. The key bit will be 0 if this integer is even, and 1 if this integer is odd. Second, Bob doesn't need to know the value of ϕ_p, so we can suppose that Alice doesn't reveal it. The protocol is certainly no less secure if the eavesdropper receives less classical information. The states Alice

sends are then seen by Bob as averaged over ϕ_p. We have

$$\rho(\phi_x, \bar{Z} = 1) = \sum_s |x = (2s + \phi_x)\sqrt{\pi}\rangle\langle x = (2s + \phi_x)\sqrt{\pi}| \quad (6.55)$$

$$\rho(\phi_x, \bar{Z} = -1) = \sum_s |x = (2s + 1 + \phi_x)\sqrt{\pi}\rangle\langle x = (2s + 1 + \phi_x)\sqrt{\pi}|. \quad (6.56)$$

Averaging over ϕ_x too and using Eq. (6.44), we see that Alice is sending a random position eigenstate. Likewise, when working with X eigenstates, Alice is sending a random momentum eigenstate. Therefore, a protocol in which Alice prepares encoded qubits and Bob measures encoded qubit operators can be replaced, without loss of security, by a simpler protocol where Alice prepares $\hat{x}$ eigenstates and $\hat{p}$ eigenstates, and where Bob performs $\hat{x}$ homodyne measurements and $\hat{p}$ homodyne measurements. Also, as discussed below, the protocol is no less secure if nonnormalisable $\hat{x}$ ($\hat{p}$) eigenstates, drawn from nonnormalisable uniform distribution, are replaced with finitely squeezed states, drawn from a broad but normalisable distribution. The conditions imposed on the necessary amount of squeezing and on these distributions for the protocol to still work will be specified below. We thus get the following protocol:

Protocol 5: BB84 with finitely squeezed states (Gottesman-Preskill) [58]. #1 Alice creates a $4(l + \delta)$-bit string b to decide for each of $4(l + \delta)$ quantum oscillator, whether it will be prepared in an x-squeezed state or in a p-squeezed state. #2 For each oscillator, she draws the value of x (or p) from a probability distribution $P_{pos}(x)$ (or $P_{mom}(p)$) #3 She sends Bob an x-squeezed state (or p-squeezed) centred on the value $(x, 0)$ (or $(0, p)$). #4 Bob receives the states and decides at random to measure them either in the x-basis or in the p-basis. #5 Alice reveals b. #6 Alice and Bob discard the oscillators for which Alice's choice of preparation and Bob's choice of measurement don't match. #7 Alice reveals $\phi_x \equiv x \mod \sqrt{\pi}$ (or $\phi_p \equiv p \mod \sqrt{\pi}$). #8 Bob subtract ϕ_x (or ϕ_p) from what he measured and adjusts the result to the nearest integer multiple of $\sqrt{\pi}$. The key bit will be 0 if this integer is even and 1 otherwise. Next steps are identical to Protocol 3.

Regular probability distributions and finite squeezing. Ideally, Protocol 5 would involve infinitely-squeezed states drawn from nonnormalisable uniform probability distributions as in Protocol 4. In fact, one can prove [58], that Protocol 5 is no less secure if we instead use squeezed states drawn from normalisable probability distributions $P_{pos(x)}$ and $P_{mom}(p)$, as long as Alice's source is exactly simulatable by measuring half of an entangled state of two oscillators. The simplest such state to think of is a two-mode squeezed state:

$$\begin{aligned} |\Psi(\Delta)\rangle_{AB} &= \frac{1}{\sqrt{\pi}} \int dx_A dx_B e^{-\frac{1}{2}\Delta^2(\frac{x_A+x_B}{2})^2} e^{-\frac{1}{2}(\frac{x_A-x_B}{2})^2/\Delta^2} |x_A, x_B\rangle \\ &= \frac{1}{\sqrt{\pi}} \int dp_A dp_B e^{-\frac{1}{2}\Delta^2(\frac{p_A-p_B}{2})^2} e^{-\frac{1}{2}(\frac{p_A+p_B}{2})^2/\Delta^2} |p_A, p_B\rangle \end{aligned} \quad (6.57)$$

If Alice measures $\hat{x}$ for her half of this state, she prepares for Bob the state

$$|\psi(x_A)\rangle = \frac{1}{(\pi\widetilde{\Delta}^2)^{1/4}} \int dx_B e^{-\frac{1}{2}(x_B - x_{B_0})^2/\widetilde{\Delta}^2} |x_B\rangle, \quad (6.58)$$

where

$$x_{B_0} = \frac{1 - \frac{1}{4}\Delta^4}{1 + \frac{1}{4}\Delta^4} x_A,$$

and

$$\widetilde{\Delta}^2 = \frac{\Delta^2}{1 + \frac{1}{4}\Delta^4}.$$

The probability distribution for the outcome of Alice's measurement reads

$$P(x_A) = \frac{\widetilde{\Delta}}{\sqrt{\pi}} e^{-\widetilde{\Delta}^2 x_A^2}. \tag{6.59}$$

Also, it is obvious from Eq. (6.57) that the probability distribution for the difference $x_A - x_B$ is governed by

$$P(x_A - x_B) = \frac{1}{\sqrt{\pi\Delta^2}} e^{-(x_A - x_B)^2/\Delta^2}. \tag{6.60}$$

So a protocol where Alice's source behaves as half of the state (6.57) is equivalent to a protocol where Alice is effectively sending Bob squeezed states with variance $\widetilde{\Delta}^2$, according to the Gaussian probability law:

$$P_{pos}(x_B) = \frac{1}{2} \int dx_A P(x_A) P(x_A - x_B), \tag{6.61}$$

and a similar expression for the probability distribution $P_{mom}(p_B)$ associated to the sending of p-squeezed states. Now, we have seen that the bit-error rate e_B is no worse than the probability that Alice's value x_A and Bob's value x_B differ by more than $\sqrt{\pi}/2$. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ denote the cp-map representing the quantum channel between Alice and Bob. We have

$$e_b < 2 \int dx_A P(x_A) \int_{\sqrt{\pi}/2}^{\infty} d\epsilon \langle x_A + \epsilon | T(|\psi)(x_A)\rangle\langle\psi)(x_A)|) | x_A + \epsilon \rangle, \tag{6.62}$$

and similarly for the phase-error rate,

$$e_p < 2 \int dp_A P(p_A) \int_{\sqrt{\pi}/2}^{\infty} d\epsilon \ \ \langle p_A + \epsilon | T\left(|\psi(p_A)\rangle\langle\psi(p_A)|\right) | p_A + \epsilon \rangle. \tag{6.63}$$

The minimum amount of squeezing required in this protocol is determined by imposing that e_b and e_p should satisfy Eq.(6.16) when T is an ideal channel. Thus we have the upper bound

$$e_b, e_p < \frac{2}{\sqrt{\pi\Delta^2}} \int_{\sqrt{\pi}/2}^{\infty} dx e^{-x^2/\Delta^2} \leq \frac{2\Delta}{\pi} e^{-\pi/4\Delta^2}. \tag{6.64}$$

In the symmetric case, $e_b = e_p$, Eq.(6.16) expresses the protocol is secure only if $\Delta < 0.784$, which corresponds to a squeezing of

$$\widetilde{\Delta} < 0.749. \tag{6.65}$$

Even though this amount of squeezing is relatively modest, we can wonder whether we cannot completely get rid of squeezing? Isn't it possible to bring a further modification transforming Protocol 5 to a secure coherent-state protocol ($\widetilde{\Delta} = 1$)?

6.5 A secure coherent-state protocol

Let us first remark that Protocol 5 was derived from an implementation of BB84 involving symmetric shift-resistant codes (Protocol 4). But the whole reasoning presented in Sect.6.4 remains valid if we start from an asymmetric code resistant to x-shifts and p-shifts satisfying Eq.(6.47). We then have

$$\Delta_x = \Delta\alpha, \quad \Delta_p = \Delta/\alpha. \tag{6.66}$$

As far as security is concerned, all that matters is that Eq.(6.16) should be satisfied. In particular, Alice could send coherent states when she chooses to use, say, the x-quadrature for encoding, at the condition of sending states having a compensating amount of squeezing when she is using the p-quadrature for encoding:

$$\Delta_x = \Delta\alpha = 1 \tag{6.67}$$

$$1 - h(e_b) - h(e_p) \geq 1 - h(\frac{2\Delta\alpha}{\pi} e^{-\pi/4(\Delta\alpha)^2}) - h(\frac{2\Delta}{\alpha\pi} e^{-\pi\alpha^2/4\Delta^2}) \geq 0. \tag{6.68}$$

Second, in Protocol 5, when Alice chooses to prepare an x-squeezed state, she draws the value of x from P_{pos} and prepares a state centred on $(x, 0)$. Similarly, when encoding with the conjugate quadrature, she prepares states centred on $(0, p)$. The decision to prepare states centred on $(x, 0)$ or $(0, p)$ relies on a convention between Alice and Bob for the axis for the x quadrature and the axis for the p quadrature. But this convention is arbitrary. For example instead of sending a state centred on $(x, 0)$, Alice could as well send a state centred on (x, p), where the key information is encoded in x, and where p is irrelevant to Bob. Thus the following protocol is as secure as Protocol 5.

Protocol 5': Modified squeezed-state encoding. #1 Alice creates a $4(l+\delta)$-bit string b to decide for each of $4(l+\delta)$ quantum oscillator, whether it will be prepared in an coherent state or in a p-squeezed state. #2 For each oscillator, she draws the value of x *and* p from probability distributions $P_{pos}(x)$, $Q_{pos}(p)$ (or $Q_{mom}(x)$, $P_{mom}(p)$). #3 She sends Bob a coherent state (or a squeezed state) centred on (x, p): $|coh(x, p)\rangle$ (or $|sq(x, p)\rangle$). #4 Bob receives the states and decides at random to measure them either in the x-basis or in the p-basis. #5 Alice reveals b. #6 Alice and Bob discard the oscillators for which Alice's choice of preparation and Bob's choice of measurement don't match. #7 Alice reveals $\phi_x \equiv x \mod \sqrt{\pi}\alpha$ (or $\phi_p \equiv p \mod \sqrt{\pi}/\alpha$). Next steps are identical to Protocol 5.

The error rates for Protocol 5' are given by:

$$\begin{aligned} e_b &= \int dx\, dp\, P_{pos}(x) Q_{pos}(p) \\ &\quad \int_{\sqrt{\pi}\alpha/2}^{\infty} de \langle x + e | T(|coh(x, p)\rangle\langle coh(x, p)|) | x + e \rangle, \end{aligned} \tag{6.69}$$

$$\begin{aligned} e_p &= \int dx\, dp\, Q_{mom}(x) P_{mom}(p) \\ &\quad \int_{\sqrt{\pi}/2\alpha}^{\infty} de \langle p + e | T(|sq(x, p)\rangle\langle sq(x, p)|) | p + e \rangle, \end{aligned} \tag{6.70}$$

where $\hat{x}|x + e\rangle = (x + e)|x + e\rangle$ and $\hat{p}|p + e\rangle = (p + e)|p + e\rangle$, and where α satisfies Eqs.(6.67)(6.68).

Now let us remark that the protocol is no less secure if Alice and Bob decide that the key is only encoded in the coherent states and never in the squeezed states. They can decide that about half of the time, Alice will send coherent states to transmit the key and to estimate e_b, while about half of the time, Alice will send squeezed states to estimate e_p. We can make a similar remark for BB84: one can decide that the key is only encoded in Z eigenstates, and that X eigenstates are only sent to determine the phase error rate. As long as e_b and e_p satisfy Eq.(6.16), the protocol will work safely. But do Alice and Bob really need to send squeezed states to estimate e_p? It seems

they don't. Since squeezed states admit a diagonal expansion in terms of coherent states [40]:

$$|sq(x,p)\rangle\langle sq(x,p)| = \int d^2\gamma\ \mathcal{P}(\gamma,x,p)|\gamma\rangle\langle\gamma|, \tag{6.71}$$

we have

$$\langle p+e|T(|sq(x,p)\rangle\langle sq(x,p)|)|p+e\rangle = \int d^2\gamma\ \mathcal{P}(\gamma,x,p)\langle p+e|T(|\gamma\rangle\langle\gamma|)|p+e\rangle, \tag{6.72}$$

i.e., e_p *can be calculated from matrix elements which can be -at least in theory* [2]*- estimated from Alice sending only coherent states and Bob measuring only the p quadrature*: coherent states allow Alice and Bob to estimate what the error rates *would have been* if Alice had sent squeezed states. This will work without, in any manner, weakening the security because if $P_{pos}, Q_{pos}, Q_{mom}, P_{mom}$ are correctly chosen, the two ensembles

$$\int dxdpP_{pos}(x)Q_{pos}(p)|coh(x,p)\rangle\langle coh(x,p)| \tag{6.73}$$

and

$$\int dxdpP_{mom}(x)Q_{mom}(p)|sq(x,p)\rangle\langle sq(x,p)| \tag{6.74}$$

are *identical*. This condition can be easily achieved if all distributions are chosen to be Gaussian. So the following protocol is equivalent to Protocol 5'.

Protocol 6: Coherent state protocol (Grosshans-Grangier)[59]. #1 Alice sends $2l$ coherent states to Bob. To determine the centre of each coherent state, she draws randomly two numbers (x,p) from a (Gaussian) probability distribution $P_{pos}(x)Q_{pos}(p)$. #2 Bob receives each state and decides randomly to measure it either in x quadrature or in p quadrature. #3 On a subset of size (about) $l/2$ of the oscillators for which Bob opted for an x-measurement, Alice and Bob estimate e_b as in Protocol 5'. On the subset of size (about) l for which Bob opted for a p-measurement, Alice and Bob estimate e_p using Eqs.(6.69)(6.72). #4 If e_b and e_p are low enough, Alice and Bob proceed with the l remaining oscillators to classical post-processing as in Protocol 5'.

6.6 More coherent state protocols

Although the discussion of the previous section shows that it is, in principle, possible to extend the Gottesman-Preskill protocol to a coherent state protocol, one could expect that such extensions will not be very efficient. According to Eqs(6.67)(6.68), the price to pay for using coherent states for x encoding in Protocol 5' implies a correspondingly low tolerance for p errors. Another reason why Protocol 6 is probably not very efficient is that the key bit assigned to a real value follows a periodic subdivision of $\mathbf{R}$ (into intervals of length $\sqrt{\pi}\alpha$ or $\sqrt{\pi}/\alpha$). On another hand, the centre of each sent coherent state is determined by drawing a Gaussian probability distribution. Hence one could expect that the assignment of bit values to real numbers should take this fact into account. Also, Protocols 4,5,5' and 6 are derived from a quantum code where *one* qubit is encoded in per oscillator. One can wonder whether it is possible to efficiently encode more qubits per oscillator, leading to more secret key bits distributed per

[2] This problem will be re-discussed in the next section.

oscillator. Finally, we would like to get free from Eq.(6.59), which seems to strongly limits the throughput information from Alice to Bob.

In the remainder of this chapter, we will derive a coherent state protocol from an entanglement purification protocol. We will first describe a protocol, where Alice and Bob extract pure EPR pairs from a tri-partite entangled state. It is precisely the use of such a state which allows to go beyond the constraint (6.59). Our purification method combines the Shor-Preskill method and a technique, called sliced error correction, allowing Alice and Bob to extract common bits from correlated real values. The principal virtue of sliced error correction is that it takes advantage of the fact that the real values sent by Alice follow a Gaussian law to encode efficiently more than one key bit per oscillator.

A secure QKD protocol with tripartite entangled states

We now describe a QKD protocol based on the use of tri-partite entangled states, which enables Alice and Bob to distill pure EPR pairs. This protocol will later be shown to be equivalent to a coherent-state protocol. Consider the following state [3]:

$$|\Psi\rangle = \int dxdp\ G(x,p)\ |x\rangle_{\mathsf{a}_1} \otimes |p\rangle_{\mathsf{a}_2} \otimes |x+ip\rangle_{\mathsf{b}} \tag{6.75}$$

where $G(x,p)$ denotes a bi-variate Gaussian distribution whose variances we need not precise now. The kets $|x\rangle$, $|p\rangle$, $|x+ip\rangle$ are shorthand notations for respectively an $\hat{x}$-quadrature eigenstate with eigenvalue x, a $\hat{p}$-quadrature eigenstate with eigenvalue p and a coherent state whose $\hat{x}$ mean value equals x and whose $\hat{p}$ mean value equals p. Note that we will from now on, and for the remainder of this chapter adopt the convention $[\hat{x},\hat{p}] = i/2$. The subscripts $\mathsf{a}_1, \mathsf{a}_2$ (resp. b) denote that the system is lying on Alice's side (resp. Bob's side). (N.B. In the following, an l-component vector $(v_1,\ldots,v_l)$ will sometimes be denoted $v_{1\ldots l}$ or $\mathbf{v}$.)

Let us first analyse a situation without eavesdropping and see how Alice and Bob can proceed to extract entangled qubits in the state $|\phi_0\rangle$ from the state $|\Psi\rangle$.

Protocol 7: QKD with tripartite entangled states

#1 Alice creates $l+t$ replicas of the state $|\Psi\rangle$ and, for each replica, she sends Bob the b part of the state. Bob confirms receipt.

#2 Alice chooses randomly t replicas that will be used for verification. She informs Bob of her choice. Using these t systems, they evaluate e_b^i and e_p^i, $i = 1\ldots m$ (see below).

#3 For each replica, Bob chooses randomly to work either with the $\hat{x}$ quadrature, either with the $\hat{p}$ quadrature and informs Alice of his choice. Let us assume, for simplicity, that Bob always chooses the $\hat{x}$-quadrature, the description of the other choice following by symmetry.

#4 Alice measures the $\hat{p}$ quadrature of the subsystem a_2, and communicates the result, say p, to Bob.

#5 Bob applies the displacement $D(0,-p)$ on the subsystem b.

#6 Alice extracts qubits from the a_1 subsystem, in applying the following linear transformation on each of the l replicas. $\mathcal{QS} : L^2(\mathbf{R}) \to L^2([0;1]) \otimes \mathcal{H}^{\otimes m}$:

$$|x\rangle \to \sigma(x)|\bar{S}(x)\rangle_\sharp \otimes |S_1(x)\rangle_{\mathsf{s}_1} \otimes \ldots \otimes |S_m(x)\rangle_{\mathsf{s}_m}. \tag{6.76}$$

The functions $\{S_i\}$ and $\bar{S}$ are defined by sliced error correction (see Appendix), the states $|\bar{s}\rangle_\sharp, 0 \le \bar{s} \le 1$ form an orthogonal basis of $L^2([0;1])$, $\sigma(x)$ is a normalisation

[3] This state has been used by Grosshans et al. independently to study the link between coherent-state QKD protocols and QKD protocols with entanglement (article in preparation).

function, and $|s_i\rangle_{\mathsf{s}_i}$ denotes an eigenstate of Z with eigenvalue $(-)^{s_i}$. By analogy with classical sliced error correction, the system s_i will be called 'slice i'.

#7 For each slice $i = 1 \dots m$, Alice and Bob agree on a CSS code that can correct le_b^i bit flips and le_p^i phase flips (if not possible due to Eq. (6.16), slice i is skipped). Alice prepares the bit-error syndrome $|\xi_i^b\rangle \in \mathcal{H}^{\otimes l_i^b}$ and the phase-error syndrome $|\xi_i^p\rangle \in \mathcal{H}^{\otimes l_i^p}$, with l_i^b (resp. l_i^p) the number of rows of H_1 (resp. $H_2^\perp$) of the i-th CSS code.

#8 Alice transmits the $2m$ syndromes $\{|\xi_i^b\rangle, |\xi_i^p\rangle\}$ to Bob as well as the $\sharp$ system.

We suppose that Alice and Bob have a noiseless quantum channel for that task. This assumption will later turn to be equivalent to the assumption of having a classical authenticated public channel available.

Now Bob extracts from the b subsystem qubits entangled with those of Alice in the state $|\phi_0\rangle$. To explain how it works, let us first rewrite, in the $\hat{x}$-basis for Bob, the state Alice and Bob share after the step #4:

$$|\psi'\rangle = \int dx dx' \; G(x,0)\gamma(x',x) \; |x\rangle_{\mathsf{a}_1} \otimes |x'\rangle_{\mathsf{b}}, \text{ where } \gamma(x',x) = \langle x'|x + i0\rangle.$$

#9 Bob applies the following linear mapping:

$$\begin{aligned}
&\mathcal{QE} : L^2([0;1])^{\otimes l} \otimes \mathcal{H}^{\otimes(l_1^b+\dots l_m^b)} \otimes L^2(\mathbf{R})^{\otimes l} \to L^2([0;1])^{\otimes 2l} \otimes \mathcal{H}^{\otimes(l_1^b+\dots+l_m^b)} \otimes \mathcal{H}^{\otimes ml} \\
&|\bar{S}(\mathbf{x})\rangle_\sharp |\xi^b_{1\dots m}\rangle |\mathbf{x}'\rangle \to \epsilon(\mathbf{x}', \bar{S}(\mathbf{x}), E_{1\dots m}, \xi^b_{1\dots m}) |\bar{S}(\mathbf{x})\rangle_\sharp \otimes |\xi^b_{1\dots m}\rangle \\
&\quad \otimes_{i=1}^m |E_i(\mathbf{x}', \bar{S}(\mathbf{x}), E_{1\dots i-1}, \xi^b_{1\dots i-1})\rangle_{\mathsf{e}_i} \otimes |\bar{E}_{m+1}(\mathbf{x}', \bar{S}(\mathbf{x}), E_{1\dots m}, \xi^b_{1\dots m})\rangle. \quad (6.77)
\end{aligned}$$

where $\epsilon(\mathbf{x}', \bar{S}(\mathbf{x}), E_{1\dots m}, \xi^b_{1\dots m})$ is a normalisation function. This mapping is explained below.

#10 By applying CSS-based EPR purification on the systems $\rho_{\mathsf{s}_i\mathsf{e}_i}$ for $i = 1 \dots m$, Alice and Bob get $\sum_i l - l_i^b - l_i^p$ pure EPR pairs in the state $|\phi_0\rangle$.

Construction of $\bar{S}$ and $\bar{E}$

First assume, for simplicity, that we have only one slice, implying the following mappings in the protocol

$$\begin{aligned}
\mathcal{QE} \circ \mathcal{QS} : |x\rangle |x'\rangle &\to \sigma(x) |S(x)\rangle \left|\bar{S}(x)\right\rangle \\
\epsilon(x', \bar{S}(x), S(x)) \left|E(x', \bar{S}(x))\right\rangle &\left|\bar{E}(x', \bar{S}(x), S(x))\right\rangle,
\end{aligned} \quad (6.78)$$

where $\sigma(x) = (d_x \bar{S}(x))^{-1/2}$ and $\epsilon(x', \bar{s}, s) = (\partial_{x'} \bar{E}(x', \bar{s}, s))^{-1/2}$, so that $\bar{S}$ and $\bar{E}$ range between 0 and 1. Thus, by linearity, a pure state $|\psi\rangle = \int dx dx' f(x,x') |x\rangle |x'\rangle$, becomes

$$(\mathcal{QE} \circ \mathcal{QS}) |\psi\rangle = \sum_{s,e \in \{0,1\}} \int d\bar{s} d\bar{e} \sigma(x) \epsilon(x', \bar{s}, s) f(x,x') |s\rangle_{\mathsf{s}} |\bar{s}\rangle_{\bar{\mathsf{s}}} |e\rangle_{\mathsf{e}} |\bar{e}\rangle_{\bar{\mathsf{e}}}, \quad (6.79)$$

where x and x' are calculated from $(s, \bar{s})$ and $(e, \bar{e}, \bar{s})$ respectively.

Our goal is to be able to extract entangled pairs in the subsystem ρ_{se} out of $|\psi\rangle$. If $\bar{S}(X)$ contains information about $S(X)$, or if $\bar{E}(X', \bar{S}(X), S(X))$ contains information about $E(X', \bar{S}(X))$, the subsystem ρ_{se} will not be pure. As an extreme example, if $S(X)$ and $E(X', \bar{S}(X))$ are perfectly correlated and if $S(X)$ can be found directly as a function of $\bar{S}(X)$, then ρ_{se} will be of the form $\rho_{\mathsf{se}} = p_0 |00\rangle \langle 00| + p_1 |11\rangle \langle 11|$, which does not allow to extract any EPR pairs.

With f real and non-negative, we can factor $\sum_{a,b\in\{0,1\}} \alpha_{ab} \left|ab\right\rangle_{\mathsf{se}}$ out of $\left|\psi\right\rangle$ by setting:

$$\sigma(x(s,\bar{s})) = \sigma_0(s) \left(f_x(x(s,\bar{s}))\right)^{-1}, \tag{6.80}$$

$$\epsilon(\bar{s}, x'(e,\bar{e},\bar{s}), s) = \epsilon_0(e,s) \left(f_{x'}(x(s,\bar{s}), x'(e,\bar{e},\bar{s}))\right)^{-1}, \tag{6.81}$$

with

$$\sigma_0^2(s) = \int_{x:S(x)=s} \left|f_x(x)\right|^2 dx, \tag{6.82}$$

$$\epsilon_0^2(e,s) = \int_{x,x':S(x)=s,E(x',\bar{S}(x))=e} \left|f_{x'}(x,x')\right|^2 dx dx', \tag{6.83}$$

and with f_x and $f_{x'}$ verifying $f(x,x') = f_x(x) f_{x'}(x,x')$, so that $\alpha_{ab} = \sigma_0(a)\epsilon_0(b,a)$. Note that $f_x(x)$ can be chosen such that $\left|f_x(x)\right|^2$ is the distribution of probability that Alice uses for modulation, and $f_{x'}(x,x')$ such that $\left|f_{x'}(x,x')\right|^2$ is the probability distribution of Bob's measured value x' conditionally to Alice sending x.

Then, we have

$$d_x \bar{S}(x) = \left|f_x(x)\right|^2 / \sigma_0^2(S(x)),$$

which means that $\bar{S}(x)$ indicates the cumulative probability

$$\bar{S}(x) = \Pr\left[X \leq x | S(X) = S(x)\right].$$

Similarly,

$$\bar{E}(x', \bar{s}, s) = \Pr\left[X' \leq x' | \bar{S}(X) = \bar{s}, S(X) = s, E(X', \bar{s}) = E(x', \bar{s})\right].$$

Each complementary function ($\bar{S}$ and $\bar{E}$) is thus chosen to have its range uniformly distributed between 0 and 1, independently of the other variables available to the party calculating it (Alice for $\bar{S}$ and Bob for $\bar{E}$).

When more than one slice is involved, this translates to:

$$\bar{S}(x) = \Pr[X \leq x | S_{1\ldots m}(X) = S_{1\ldots m}(x)], \tag{6.84}$$

$$\begin{aligned} \bar{E}_{m+1}(x', \bar{s}, s_{1\ldots m}) = \Pr[X' \leq x' | \bar{S}(X) = \bar{s} \\ \wedge\ S_{1\ldots m}(X) = s_{1\ldots m} \\ \wedge\ E_1(X', \bar{s}) = E_1(x', \bar{s}) \wedge \ldots \\ \wedge\ E_m(X', \bar{s}, s_{1\ldots m-1}) = E_m(x', \bar{s}, s_{1\ldots m-1})]. \end{aligned} \tag{6.85}$$

From Protocol 7 to a coherent state protocol

It is easy to understand that Protocol 7 is equivalent to a coherent-state protocol. First, we note that Protocol 7 is unaffected if Alice applies the mapping $\mathcal{QS}$ prior to sending Bob the b part of the state (6.75). In addition, if Alice and Bob are confident that they can perform entanglement purification with the pairs $\mathsf{s}_i\mathsf{e}_i$, they do not need to correct phase errors. They will anyway extract only (classical) bits from Z measurements performed on these qubits. Hence, there is no need for Alice to measure phase-error syndromes for each slice. Then, instead of measuring the syndromes $\left|\xi_i^b\right\rangle$, transmitting them to Bob and measuring the encoded qubits, Alice can measure Z on each qubit $\left|S_i(x_j)\right\rangle$, $i = 1\ldots m, j = 1\ldots l$, and send Bob the *classical* syndromes ξ_i^b. To further simplify the protocol, Alice can as well measure $\hat{x}_{\mathsf{a}_1}$ at step #2, compute the syndromes ξ_i^b and send them to Bob. If both the $\hat{x}_{\mathsf{a}_1}$ quadrature and the $\hat{p}_{\mathsf{a}_2}$

quadrature are measured at step #1, there is no need for Alice to bother preparing such a complicated state as Eq. (6.75). The protocol is no less secure if Alice sends Bob a coherent state whose $\hat{x}$ mean value and $\hat{p}$ mean value are drawn randomly from the distribution $|G(x,p)|^2$. A similar argument shows that there is no need for Bob to perform the sequence of complicated operations $\Omega\mathcal{E}_i$, perform both bit and phase error correction, and measure the encoded qubits. He can as well measure the $\hat{x}$ quadrature on the b subsystem, apply classical sliced error correction and privacy amplification for each slice. It also becomes useless for Alice to send the mean value of $\hat{p}_{\mathsf{a}_2}$.

We have thus reduced the protocol described above to a much simpler protocol. However, it is still essential for Alice and Bob to have an estimation of the error rate e_b^i and e_p^i, for each slice. This problem is examined below. Let us summarise the reduced protocol.

Protocol 8: Alternative coherent state protocol #1 Alice draws randomly two real numbers x and p according to the Gaussian distribution $|G(x,p)|^2$, prepares a coherent state $|x+ip\rangle$, and sends it to Bob over the quantum channel. #2 Bob receives the coherent state, and decides randomly to measure either $\hat{x}$, or $\hat{p}$. Let us assume he measures $\hat{x}$. #3 Alice and Bob repeat these two first steps $(l+t)$ times, except that with t randomly chosen samples, Bob makes a homodyne detection in a randomly chosen direction θ. This will allow Alice and Bob to estimate e_b^i and e_p^i (see below). #4 Alice and Bob apply (classical) sliced error correction. The binary error correction protocol used consists of sending the syndromes of the error correcting code C_1^i used in Protocol 7, where C_1^i refers to the CSS code used for slice i. #5 Let k_i denote the sifted key obtained for each slice. Each slice i provides a private key $k_i + C_2^i$ (privacy amplification), where $C_2^i \subset C_1^i$ refers to the CSS code used for slice i.

Bit error rates and phase error rates

If the joint state of Alice and Bob ρ_{ab} is known, the bit error rate and phase error rate can be easily computed. Let $\rho_{\mathsf{s}_i\mathsf{e}_i} = \mathrm{Tr}_{\mathrm{All}\setminus\mathsf{s}_i\mathsf{e}_i}(\rho_{\mathsf{ab}})$. Then, the bit error rate is $e_b^i = \mathrm{Tr}(\rho_{\mathsf{s}_i\mathsf{e}_i}(|\phi_2\rangle\langle\phi_2|+|\phi_3\rangle\langle\phi_3|))$ and the phase error rate $e_p^i = \mathrm{Tr}(\rho_{\mathsf{s}_i\mathsf{e}_i}(|\phi_1\rangle\langle\phi_1|+|\phi_3\rangle\langle\phi_3|))$.

The bit error rate e_b^i is easy to estimate. We have:

$$e_b^i = \Pr[S_i(X) \neq E_i(X', \bar{S}(X), S_{1\ldots i-1}(X))],$$

and these quantities can be estimated with high statistical confidence by Alice sending coherent states, and Bob performing x quadrature measurements. The phase error rate, however, is more difficult to estimate. Still, it is possible – at least in theory – to evaluate it.

One way is to fully characterise the quantum channel T between Alice and Bob. This can be achieved if we know $T(|x'\rangle\langle x''|)$, for all $x', x'' \in \mathbf{R}$. By sending a coherent state $|x+ip\rangle_{\mathsf{b},\mathrm{coh}}$ to Bob, he can estimate $T(|x'\rangle\langle x''|)$, using homodyne measurements in all quadratures.

$$T(|x+ip\rangle\langle x+ip|) \propto \int dx'dx'' e^{-(x'-x)^2/4N_0-(x''-x)^2/4N_0} e^{i(x'-x'')p/2N_0} T(|x'\rangle\langle x''|). \tag{6.86}$$

By setting $D = x' - x''$ and $S = x' + x'' - 2x$, we get

$$T(|x+ip\rangle\langle x+ip|) \propto \int dDdS e^{-S^2/8N_0-D^2/8N_0+iDp/2N_0} T(|x+S+D\rangle\langle x+S-D|), \tag{6.87}$$

which shows that we can get the knowledge of $T(|x+ip\rangle\langle x+ip|)$ yields $T(|x'\rangle\langle x''|)$ integrated with an invertible kernel (Gaussian convolution in S, multiplication by $e^{-D^2/8N_0}$ and Fourier-transform in D).

With Alice sending many coherent states, and Bob performing quadrature measurements in all directions, it is in principle possible to determine the operators $T(|x+ip\rangle\langle x+ip|)$ (quantum tomography), and thus to deduce $T(|x'\rangle\langle x''|)$. In practice, however, this can be a difficult task. First, the complete re-construction of the density matrices $T(|x+ip\rangle\langle x+ip|)$ for many x and p may require an unacceptably large number of samples. Second, the inversion of the integration to find $T(|x'\rangle\langle x''|)$ is unlikely to be accurate.

To address these problems, we propose the following two ideas. First, Alice and Bob can agree on a modelled channel, parametrised by only a few variables $p_1, \ldots, p_n$ (e.g., losses, added noise and possibly some non-Gaussian effect), which best suits the reality of their apparati. Instead of performing a complete non-parametric re-construction, they only need to estimate $p_1, \ldots, p_n$. To ensure that the channel actually follows the estimated model, a statistical hypothesis test is done after the estimation. If this test fails, Alice and Bob shall either abort the protocol or agree on a better channel model.

Second, the multiplication by $e^{-D^2/8N_0}$ shows that the terms involving $|x'\rangle_{\mathsf{b}}\langle x''|$ for distant x' and x'' (i.e., $|x'-x''| \gg 1$) are difficult to estimate with high accuracy. Therefore, it is difficult to estimate the phase coherence of the slices that involve distant values (e.g., the most significant bit in the numerical example below). To avoid such problems, one can simply make such a slice public by integrating it into the definition of $\bar{S}(x)$ since its secrecy is difficult to quantify.

The Attenuation Channel

The attenuation channel can be modelled as if Eve installs a beam-splitter in between two sections of a lossless line, sending vacuum at the second input.

We assume that Alice sends coherent states with a modulation variance of $31\times$ vacuum noise, which gives Alice and Bob up to $I(A;B) = 2.5$ common bits in absence of losses or noise. This matches the order of magnitude implemented in [10].

First, let us investigate a simple case with only one slice and with a slice estimator that does not depend on $\bar{S}(x)$. The mapping is the following:

$$|x\rangle_{\mathsf{a}_1}|x'\rangle_{\mathsf{b}} \to \sigma(x)\epsilon(x')\,|S(x)\rangle_{\mathsf{s}}\left|\bar{S}(x)\right\rangle_{\bar{\mathsf{s}}}|E(x')\rangle_{\mathsf{e}}\left|\bar{E}(x')\right\rangle_{\bar{\mathsf{e}}}, \tag{6.88}$$

with $S(x) = 0$ when $x \leq 0$ and $S(x) = 1$ otherwise, $E(x') = 0$ when $x \leq 0$ and $E(x') = 1$ otherwise, $\bar{S}(x) = \Pr[X \leq x | S(X) = S(x)]$ and $\bar{E}(x') = \Pr[X' \leq x' | E(X') = E(x')]$.

When the entangled state $|\Psi\rangle$ is pure, that is without any eavesdropping, the substate ρ_{se}, obtained by tracing out everything but s and e, is numerically calculated. This state has a bit error rate $e_b = 5.65\%$ and a phase error rate $e_p = 8.73\%$, which makes it possible to extract $R = 1 - h(e_b) - h(e_p) \approx 0.259$ secret bits per sample. Adding a small attenuation of 0.05 dB to the channel, we get $e_b = 5.68\%$, $e_p = 12.9\%$ and $R \approx 0.131$. The rate R drops to 0.037 for a 0.1 dB attenuation, and it is not possible to go much further.

Using the full construction with two slices, we were able to get the EPR rates described in Fig. 6.1. The slices S_1 and S_2 are defined by dividing the real axis into four equiprobable intervals labelled by two bits. S_1 represents the least significant bit, and S_2 the most significant. For the case with low losses, it is thus possible to distill more than one EPR pair per sample. Also, note that the phase error rate increases faster with the attenuation for ρ_2 than for ρ_1, with $\rho_i = \rho_{\mathsf{s}_i\mathsf{e}_i}$. This intuitively follows from the fact that the information Eve can gain from her output of the beam splitter affects first the most significant bit contained in $S_2(x)$.

Due to the higher bit error rate in ρ_1, it was not possible to distill EPR pairs in slice 1 with losses beyond 0.7 dB. It was however still possible to distill EPR pairs in slice 2, up to 1.4 dB losses (about 10km with fiber optics with losses of 0.15db/km).

	ρ_1			ρ_2		
Losses	e_b	e_p	R	e_b	e_p	R
0.0dB	3.11%	5.33%	0.752	0.0000401	0.710%	0.938
0.4dB	3.77%	13.7%	0.193	0.0000782	28.6%	0.135
0.7dB	4.32%	20.0%	0.0204	0.000125	37.5%	0.0434
1.0dB	-			0.000194	42.3%	0.0147
1.4dB	-			0.000335	45.6%	0.00114

Figure 6.1: Error rates and secret key rates with two slices in an attenuation channel.

This result does not pose any fundamental limit, as it can vary with the modulation variance and with the choice of the functions S_1 and S_2.

Finally, note that although this example involves a Gaussian channel, this particularity is not exploited here and such a calculation can be as easily done for a non-Gaussian and/or collective attack (modulo the possible difficulty to calculate phase errors).

6.7 Summary and discussion

We have reviewed the Shor-Preskill proof of security of the BB84 protocol, the construction of shift resistant codes, and how the two can be combined to devise a secure squeezed protocol: the Gottesman-Preskill protocol [58]. This construction leads to protocols robust against any eavesdropping strategy, including collective attacks. In contrast, the Cerf-Lévy-Van Assche [53, 54] construction addresses only Gaussian individual attacks. Nevertheless, both constructions lead to a protocol where squeezed states are modulated according to a Gaussian law, and where the indistinguishability of two different mixtures of squeezed states plays a crucial role. It is tempting to conjecture that Gaussian individual attacks should be ultimately optimal for eavesdropping.

We have then showed how the Gottesman-Preskill protocol can be extended into a coherent state protocol, whose physical part is identical to the protocol of Grosshans and Grangier [66]. From a fundamental point of view, this extension is very important, because it shows that, in principle, no non-classical feature of light, such as squeezing, is required for secure quantum key distribution.

Then, we have studied the possibility to combine entanglement purification protocols with sliced error correction. Our aim was to establish secure coherent state protocols where the assignment of bits to a real value doesn't follow a periodic subdivision of $\mathbf{R}$ as in the Gottesman-Preskill protocol. We have shown that key distribution was possible up to 1.4 dB (for a modulation on Alice's side of 31× vacuum noise). This value should be compared to the limit imposed by the symmetric cloning attack, which has been shown to be equivalent to a 3 dB loss in [59]. Although, we have considered a Gaussian individual attack (equivalent to a cloning machine) as in Chapter 5, the approach developed here is different. In Chapter 5, one explicitly constructs an eavesdropping strategy and derives a noise level *above* which the protocol is unsecure: 3 dB. Here, we don't make any hypothesis a priori, we establish general security conditions, and *then*, considering a specific attack, we establish a noise level below which the protocol can be made secure: 1.4 dB (for a modulation on Alice's side of 31× vacuum noise). This 1.4 dB threshold relies on the design of our code, i.e., our design of slice and estimator functions, for which we have no proof of optimality. This leaves open the possibility of finding better codes robust at higher noise levels.

6.8 Appendix: Sliced error correction

We here recall the main principles of sliced error correction (SEC) in a form that is slightly different from the presentation in [67]. To suit our needs, we here describe SEC in terms of *invertible* functions giving the slices and the estimators, and error correction is operated by sending syndromes of classical linear error-correcting codes (ECC) as binary correction protocols. Furthermore, we explicitly restrict ourselves to the case of scalar values.

Suppose Alice and Bob have l couples of correlated random real variables:

$$(x_1, x'_1), \dots, (x_l, x'_l).$$

Sliced error correction aims at providing Alice and Bob with a means to extract m common bits: $S_1(x), \dots S_m(x)$.

First, let us describe how Alice the m bits: $S_1(x), \dots S_m(x)$ from a real number x. To make the mapping invertible, she also needs the function $\bar{S}(x)$ such that the set $\{\bar{S}(x), S_{1\dots m}(x)\}$ allows to recover x, for all $x \in \mathbf{R}$. Formally, we define $\mathcal{S} : \mathbf{R} \to [0;1] \times \mathbf{F}_2^m : x \to (\bar{S}(x), S_{1\dots m}(x))$.

To make things more concrete, the functions $S_i(x)$ cut the real line into intervals (see [67] for more details), whereas $\bar{S}(x)$ indicates some value that allows one to find x within a given interval.

Then, for each bit vector ("slice") $S_i(x_{1\dots l}) = (S_i(x_1), \dots S_i(x_l))$, Alice sends $\bar{S}(x_{1\dots l})$ together with the syndrome $\xi_i = H_i S_i(x_{1\dots l})$ to Bob, where H_i is the $l'_i \times l$ parity check matrix of an ECC.

Bob also converts his variables x' into bits and wishes to estimate the best he can the bits $S_i(x_{1\dots l})$. In addition to that, he waits to have enough information for correcting the bit vector i before converting x' into a bit vector that estimates $S_j(x)$, $j > i$. To estimate the first slice $S_1(x_{1\dots l})$, he uses a function $E_1(x', \bar{S}(x))$ that gives him the best estimate of $S_1(x)$ using the knowledge of x' and $\bar{S}(x)$. From $E_1(x'_{1\dots l}, \bar{S}(x_{1\dots l}))$ and ξ_1, we assume he has enough information to recover $S_1(x_{1\dots l})$ with high probability. Then, for $i > 1$, Bob estimates $S_i(x_{1\dots l})$ using the estimator $E_i(x'_{1\dots l}, \bar{S}(x_{1\dots l}), E_{1\dots i-1}(x'_{1\dots l}, \dots), \xi_{1\dots i-1})$.

Note that we can also write an estimator E_i as $E_i(x', \bar{S}(x), S_{1\dots i-1}(x))$. Even though it is an improper notation, since Bob does not have access to x nor to $S_{1\dots i-1}(x)$, this simplified notation is sensible since Bob has enough information to recover $S_{1\dots i-1}(x)$.

We add an extra function $\bar{E}_{m+1}(x', \bar{S}(x), S_{1\dots m}(x))$, which is required to make the mapping $\mathcal{E}$ defined below invertible. Bob's mapping thus formally writes:

$$\begin{aligned}
&\mathcal{E} : [0;1]^l \times \mathbf{F}_2^{l'_1+\dots+l'_m} \times \mathbf{R}^l \to [0;1]^{2l} \times \mathbf{F}_2^{l'_1+\dots+l'_m} \times \mathbf{F}_2^{lm} : \\
&(\bar{\mathbf{s}}, \xi_{1\dots m}, \mathbf{x}') \to (\bar{\mathbf{s}}, \xi_{1\dots m}, E_1(\mathbf{x}', \bar{\mathbf{s}}), \dots, \\
&\quad E_m(\mathbf{x}', \bar{\mathbf{s}}, E_{1\dots m-1}(\mathbf{x}', \dots), \xi_{1\dots m-1}), \bar{E}_{m+1}(\mathbf{x}', \bar{\mathbf{s}}, E_{1\dots m}(\mathbf{x}', \dots), \xi_{1\dots m})). \quad (6.89)
\end{aligned}$$

At the end, Bob has enough information to recover the $m \times l$ bits $S_{1\dots m}(x_{1\dots l})$, out of which $\sum_i l'_i$ parity bits were revealed.

Chapter 7

Conclusions and Perspectives

Quantum information cannot be cloned. In this thesis, we have contributed to give a quantitative meaning to this fundamental principle. Focusing mainly on continuous variable systems, we have studied approximate quantum cloning machines and quantum key distribution protocols.

We have studied how the quantum information contained in one or more quantum system distributes amongst the clones. We have seen that the information that can be extracted from the clones depends on the manner it is encoded in the input.

In quantum key distribution, classical information sent by an emitter is distributed between an authorised receiver and a potential eavesdropper. We have analysed the security of protocols where the emitter only sends coherent states of light, and the receiver only performs homodyne measurements. Such protocols have emerged recently as a promising alternative to other quantum key distribution protocols based on the use of non-classical states of light, difficult to prepare, such as squeezed states or Fock states. Our study takes root at a beautiful interplay between quantum key distribution and quantum error correcting codes.

Future work will be mainly concerned with the secrecy capacity and the quantum capacity of a quantum channel. The quantum capacity of a quantum channel is a quantity defined similarly to the classical capacity: it is the (asymptotically) maximal number of qubits that one can transmit over this channel per use of the channel. To date, we don't know the quantum capacity of any single channel. Concentrating on Gaussian channels, we plan to develop explicit quantum error correcting codes and provide achievable rates over such channels.

The secrecy capacity of a quantum channel is the (asymptotically) maximal number of secret bits that one can transmit per use of this channel. The quantum capacity trivially yields a tight lower bound on the secrecy capacity, for a quantum code allows to purify noisy EPR pairs, and one secret key bit can be extracted (locally) by two parties sharing an EPR pair. However, it is not known whether there are channels for which the secrecy capacity is strictly higher than the quantum capacity. Again focusing on Gaussian channels, we would like to further develop coding theory, and provide achievable secrecy rates.

In conclusion, we have tackled the issue of distributing the information contained in a quantum state by studying two classes of problems: quantum cloning and quantum cryptography.

Let us start with quantum cloning. We have studied the cloning of orthogonal qubits and identified a class of situations where orthogonal pairs of qubits contain more information than a non-orthogonal one. It is tempting to conclude that orthogonal pair of qubits always contain more information than non-orthogonal ones. But it is not

quite so. In fact, the optimal kind of quantum information carriers is strongly related to the kind of information one wants to extract from it [41]. At the time being, we don't know much about such relations. To solve this problem in general is probably too big a problem. However, a future line of research might be to provide the basis of a systematic study of such relations.

We have also provided a detailed analysis of optimal quantum cloning transformation. If the cloner is required to be covariant with respect to translation and rotation in phase-space, and if optimality is measured by the noise introduced by cloning, then we have seen that Gaussian cloners are optimal. However, similarly to the situation of qubits, such a cloner is no more optimal for another figure of merit.

Finally, we have studied the security of coherent-state quantum key distribution protocols. We have provided a means to assess the security of such a protocol under assumptions which are more general than the Gaussian attack. We have seen how this issue is connected to that of determining the quantum capacity and the secrecy capacity of a continuous quantum channel, which shows once again the importance of these problems. An important result is to have designed quantum codes exhibiting features of classical sliced error correction. We plan to develop further such codes and hope to find achievable rates for quantum codes over continuous channels in general, and over Gaussian channels in particular.

Bibliography

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 623 (1948).

[2] *Information Theory, 50 years of discovery*, edited by S. Verdú and S. W. M. Laughlin (IEEE Press, New York, 2000).

[3] G. Alber, T. Beth, M. Horodecki, and P. Horodecki, *Quantum Information. An Introduction to Basic Theoretical Concepts and Experiments* (Springer-Verlag, Berlin Heidelberg, 2001).

[4] P. W. Shor, in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, New York, 1994), pp. 124–134, e-print quant-ph/9508027.

[5] R. Rivest, A. Shamir, and L. Adelman, MIT Laboratory for Computer Science (1979), technical Report MIT/LCS/TR-212.

[6] R. Somma *et al.*, Phys. Rev. A **65**, 042323 (2002).

[7] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[8] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1865 (1993).

[9] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[10] F. Grosshans *et al.*, Nature **421**, 238 (2003).

[11] S. Haroche, Physics Today **49**, 51 (1996).

[12] I. Marcikic *et al.*, Nature **421**, 509 (2003).

[13] N. J. Cerf and S. Iblisdir, Phys. Rev. A **62**, 040301 (2000).

[14] S. L. Braunstein *et al.*, Phys. Rev. Lett. **86**, 4438 (2001).

[15] N. J. Cerf and S. Iblisdir, Phys. Rev. A **64**, 032307 (2001).

[16] N. J. Cerf and S. Iblisdir, Phys. Rev. Lett. **87**, 247903 (2001).

[17] J. Fiurasek, S. Iblisdir, S. Massar, and N. J. Cerf, Phys. Rev. A **65**, 040302(R) (2002).

[18] R. F. Werner, Phys. Rev. A **58**, 1827 (1998).

[19] Proceedings of the American Mathematical Society 211 (1955).

[20] M. Keyl, Phys. Rep. **5**, 431 (2002).

[21] J. Preskill, (1998), lecture Notes for Physics 229: Quantum Information and Computation.

[22] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[23] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).

[24] D. Dieks, Phys. Lett. A **92**, 271 (1982).

[25] R. Laflamme, C. Miquel, J. Paz, and W. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[26] V. Buzek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).

[27] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).

[28] D. Bruß *et al.*, Phys. Rev. A **57**, 2368 (1997).

[29] D. Bruß, A. Ekert, and C. Macchiavello, Phys. Rev. Lett **81**, 2598 (1998).

[30] M. Keyl and R. F. Werner, Ann. H. Poincaré **2**, 1 (2001).

[31] N. J. Cerf, Phys. Rev. Lett. **84**, 4497 (2000).

[32] N. J. Cerf, J. Mod. Opt. **47**, 187 (2000).

[33] N. Gisin and S. Popescu, Phys. Rev. Lett. **83**, 432 (1999).

[34] S. Iblisdir, J. Fiurasek, S. Massar, and N. Cerf, Proceedings of the Sixth International Conference on Quantum Communication, Measurement and Computing To appear (2002).

[35] C. Simon, G. Weihs, and A. Zeilinger, Phys. Rev. Lett. **84**, 2993 (2000).

[36] V. Buzek and M. Hillery, Phys. Rev. Lett **81**, 5003 (1998).

[37] J. Fiurasek, Phys. Rev. A **64**, 062310 (2001).

[38] N. J. Vilenkin and A. Klimik, *Representations of Lie Groups and Special Functions* (Kluwer, Dordrecht, Netherlands, 1991).

[39] K. Audenaert and B. D. Moor, Phys. Rev. A **65**, 030302 (2002).

[40] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer-Verlag, Berlin, 1994).

[41] S. Massar, Phys. Rev. A **62**, 040101 (2000).

[42] E. Arthurs and J. L. Kelly, Jr, Bell Syst. Tech. J. **44**, 725 (1965).

[43] A. Holevo, *Probabilistic and Statistical Aspects of quantum theory* (North-Holland, Amsterdam, 1982).

[44] C. Helstrom, *Quantum detection and estimation theory* (Academic Press, New York, 1976).

[45] N. J. Cerf, A. Ipe, and X. Rottenberg, Phys. Rev. Lett. **85**, 1754 (2000).

[46] C. M. Caves, Phys. Rev. D **26**, 1817 (1982).

[47] P. Kok, H. Lee, and J. P. Dowling, Phys. Rev. A **65**, 052104 (2002).

[48] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001).

[49] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[50] S. Stenholm, Annals of Physics **218**, 233 (1992).

[51] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997).

[52] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.

[53] N. J. Cerf, M. Lévy, and G. Van Assche, Phys. Rev. A **63**, 052311 (2001).

[54] N. J. Cerf, S. Iblisdir, and G. Van Assche, Eur. Phys. J. D **18**, 211 (2002).

[55] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley & Sons, New York, 1991).

[56] T. C. Ralph, Phys. Rev. A **62**, 062306 (2000).

[57] U. M. Maurer, IEEE Trans. Inform. Theory **39**, 733 (1993).

[58] D. Gottesman and J. Preskill, Phys. Rev. A **63**, 022309 (2001).

[59] F. Grosshans and P. Grangier, Phys. Rev. Lett. **88**, 057902 (2002).

[60] A. Beveratos *et al.*, Phys. Rev. Lett. **89**, 187901 (2002).

[61] M. Koashi and J. Preskill, Phys. Rev. Lett. **90**, 057902 (2003).

[62] A. R. Calderbank and P. Shor, Phys. Rev. A **54**, 1098 (1996).

[63] A. Steane, Proc. R. Soc. London A **452**, 2551 (1996).

[64] H. K. Lo and H. Chau, Science **283**, 2050 (1999).

[65] D. Gottesman, A. Kitaev, and J. Preskill, Phys. Rev. A **64**, 012310 (2001).

[66] F. Grosshans and P. Grangier, Phys. Rev. A **64**, 010301 (2001).

[67] G. Van Assche, J. Cardinal, and N. J. Cerf, arXiv e-print cs.CR/0107030 (unpublished).